

## STIC-Biotech/ChemLib

From:

Sent:

Schnizer, Richard Wednesday, August 27, 2003 2:28 PM STIC-Biotech/ChemLib 09/965,594

To:

Subject:

Please search the commercial and published application databases for polypeptide SEQ ID NOS: 1, 12, 14, 16, 18, 20, 22, and 26 from 09/965,594. Please also search for nucleic acids that could encode these polypeptides.

Thank you-

Richard Schnizer, Ph.D. Patent Examiner Art Unit 1635 CM1 12E17 703-306-5441 Mail Box CM1 11E12

Searcher:
Phone:
Location:
Date Picked Up:
Date Completed:
Searcher Prep/Review:
Clerical:
Online time:

TYPE OF SEARCH:
NA Sequences:
AA Sequences:
Structures:
Bibliographic:
Litigation:
Full text:
Patent Family:

vendor/cost (w	here applic.)
STN:	
DIALOG:	
Questel/Orbit:	
DRLink:	
Lexis/Nexis:	
Sequence Sys.:	
WWW/Internet:	
Other (specify)	

```
AR11866 Sequence 109331 Sequence 109331 Sequence 109331 Sequence 1871872 Sequence AR11872 Sequence AR11872 Sequence 108294 Sequence 108294 Sequence 176483 Sequence 176484 Sequence 176484 Sequence 176484 Sequence 176485 Sequence 176485 Sequence 176486 Seq
                                                                                                                                                                                                                                                                                                                                                                                                                     AF369218 Hepatitis
AF369235 Hepatitis
AR118686 Sequence
106434 Sequence 48
109328 Sequence 8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              AF369224 Hepatitis
AF369237 Hepatitis
AF511948 Hepatitis
AF369230 Hepatitis
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 E08263 gRNA of Hep
E08264 cDNA of Hep
D10749 Hepatitis C
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        AR118692 Sequence
106440 Sequence 54
IO9329 Sequence 10
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             AX441176 Sequence
AX467113 Sequence
                                                                                                                                                                                                                                                                                                                                                                                      AX395309 Sequence AX454818 Sequence
                                                                                                                                                                                                                                    Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.
                                                                                                                                                                                                                                                                                                                                                      Description
                                                                                                                                                                                                                                                                                                       SUMMARIES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           HPCPLYPRE
AF387805
                                                                                                                                                                                                                                                                                                                                                                                                                       AF369218
AF369235
AR118686
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       AR118696
I09331
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            HPCPOLYP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        AR118722
AR118723
BD091382
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           AR301300
AR176483
BD080334
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                AF387806
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                AF369224
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           AR166930
                                                                                                                                                                   em_htgo_hum:*
em_htgo_mus:*
em_htgo_other:*
                                                  em_htg_other:*
                   em_htg_hum:*
                                                                      htg_mus:*
                                                                                  em_htg_pln:*
em_htg_rod:*
                                                                                                                     em_htg_mam:*
em_htg_vrt:*
                                                                                                                                                                                                                                                                                                                                                        DB
                                                                                                                                                                                                                                                                                                                     %
Query
Match Length D
                                                                                                                                                                                                                                                                                                                                                                                                                                       543
5360
5360
5360
6785
6785
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        9185
9185
9185
9185
9379
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           9401
9609
9609
9618
9646
9693
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           9401
                                                                                                                                                                                                                                                                                                                                                                                                    Score
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       RESULI 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                        Š
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             U
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Command line parameters:
-MODEL-frame+_pcn.model -DEV=xlp
-MODEL-frame+_pcn.model -DEV=xlp
-MODEL-frame+_pcn.model -DEV=xlp
-Cgn2_1/USFT0_spool/US0995594/runat_29082003_151919_28310/app_query.fasta_i.2872
-OB-cgn2_1/USFT0_spool/US0995594/runat_29082003_151919_28310/app_query.fasta_i.2872
-DB-cenbmbl .OFMT=fastap -SUFFIX=+ge -MINMATCH=0.1 -LCOPCL=0 -LCOPEXT=0
-UNTS-bits -START=i -EMD=-i -MATRIX=blosum62 -TRANS-human40.cdi -LIST-45
-UNCALIGH=200 -THF_SCORE_pct -THF_MAXIEN-0 -MAXIEN-2000000000
-USFRT=USPECORE_pct -THF_AF666_@WINDLEN-0 -MAXIEN-2000000000
-USFRT=USPECORE_PCT -MATRIX=DS-PELORE-100 -LONGLOG
-DEV_TIMBOUT=120 -WARR_TIMBOUT=30 -THREADS=1 -XGAPOP=10 -XGAPDET=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7
                                                                                                              August 30, 2003, 19:18:33; Search time 2365.6 Seconds (without alignments) 3147.423 Million cell updates/sec
                                                                                                                                                                                                 953
1 MAPITAYAQQTRGLLGCIIT......GVAKAVDFIPVESLETIMRS
              GenCore version 5.1.6
Copyright (c) 1993 · 2003 Compugen Ltd.
                                                                               - nucleic search, using frame_plus_p2n model
                                                                                                                                                                                                                                                                                                                                                                               Total number of hits satisfying chosen parameters:
                                                                                                                                                                                                                                                                                                                                                   2888711 segs, 20454813386 residues
                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Post-processing: Minjmum Match 0%
Maximum Match 100%
Listing first 45 summaries
                                                                                                                                                                                                                                                                  C.5
0.5
7.0
                                                                                                                                                                                                                                                            Xgapop 10.0 , Xgapext Ygapop 10.0 , Ygapext Fgapop 6.0 , Fgapext Delop 6.0 , Delext
                                                                                                                                                                                                                                                                                                                                                                                                                    Minimum DB seq length: 0 Maximum DB seq length: 20000000000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       GenEmbl:*

1: 90_ba:*

2: 90_htg:*

3: 90_btg:*

1: 90_ow:*

90_ov:*

90_ph:*

90_ph:*

90_ph:*

90_ph:*

90_ph:*
                                                                                                                                                                                    US-09-965-594-1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        9b_sts: *
9b_sy: *
9b_un: *
9b_vi: *
em_ba: *
em_fun: *
em_fun: *
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            em_in:*
em_mu:*
em_on:*
em_or:*
em_pat:*
em_ph:*
em_pl:*
em_pl:*
em_pl:*
                                                                                                                                                                                                                                                    BLOSUM62
                                                                                                                                                                                    Title:
Perfect score:
                                                                                                                                                                                                                                                    Scoring table:
                                                                                  OM protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Database :
                                                                                                                                                                                                                      Sequence:
                                                                                                                                                                                                                                                                                                                                                    Searched:
                                                                                                                   ou:
```

~

160 480

```
/transl_table=11
/transl_table=11
/transl_table=11
/protein_id="121714048"
/db_xref="61"=121714048"
/db_xref="61"=1217140048"
/dc_xref="61"=1217140048"
/dc_xre
                                                                                                                                                                                                                                                                                                                                                                                                          PAT 06-JUL-2002
121 GlyAspSerArgGlySerLeuLeuSerProArgProIleSerTyrLeuLysGlySerSer
                                                                                           /note-"representative NS3/4a conformational antigen
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ., Tandeske, L., George-Nasciemento, C.,
                                                                                                                                                                                                                                                                                                                                                                                                          linear
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       2058
180
2
0
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          // >2058
/note-"unnamed protein product"
/codon_start-1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  /organism="synthetic construct"
/mol_type="genomic DNA"
/db_xref="taxon:32630"
                                                                                                                                                                                                                                                                                                                                                                                                            DNA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Immunoassays for anti-hcv antibodies
Patent: WO 0196870-A 1 20-DEC-2001;
CHIRON CORPORATION (US)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Length:
Matches:
                                                                                                                                                                                                                                                                                                                                                                                                 Sequence 1 from Patent W00196870.
AX454818.1 GI:21714047
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Gaps:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            US-09-965-594-1 (1-182) x AX454818 (1-2058)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Location/Qualifiers
1. .2058
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Chien, D.Y., Arcangel, P., Tan
Coit, D. and Medina-Selby, A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    581
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           6.56e-68
946.00
100.00%
98.90%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       artificial sequences.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    synthetic construct
synthetic construct
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  633 c
                                                                                                                                                                                                                                                                                               181 ArgSer 182
                                                                                                                                                                                                                                                                                                                                  541 AGGTCC 546
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Percent Similarity:
Best Local Similarity:
Query Match:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    419
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Alignment Scores:
 01
                                     301
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        source
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  BASE COUNT
ORIGIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Pred. No.:
                                                                                                                                                                                                                                                                                                                                                                                                                            DEFINITION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ORGANISM
                                                                                                                                                                                                                                                                                                                                                                                                                                                ACCESSION
VERSION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              TITLE
JOURNAL
                                                                                                                                                                                                                                                                                                                                                                     RESULT 2
AX454818
LOCUS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 KEYWORDS
SOURCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           REFERENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             AUTHORS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CDS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     FEATURES
                                   a
                                                                                                       a
                                                                                                                                                õ
                                                                                                                                                                               g
                                                                                                                                                                                                                       οy
                                                                                                                                                                                                                                                          G
                                                                                                                                                                                                                                                                                               ó
                                                                                                                                                                                                                                                                                                                                  q
                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                          PAT 18-MAY-2002
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  50
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            4 C
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              MetalaProlleThrAlaTyrAlaGlnGlnThrArgGlyLeuEcuGlyCyslie:leThr
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     21 SerLeuThrGlyArgAspLysAsnGlnValGluGlyGluValGlnIleValSerThrAla
                                                                                                                                                                                                                                                                                                                                                           /note="representative NS3/4a conformational antigen"
1. .>2058
                                                                                                                                                                                 Chien, D.Y., Arcangel. P., Tandeske, L., George-Nasciemento, C. Coit, D. and Mcdina-Selby, A.

For antigen/antibody combination assay
Patent: WO 0196875-A 2 20-DEC-2001;
CHIRON CORPORATION (US)
                     linear
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               2058
180
2
0
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Length:
Matches:
Conservative:
Mismatches:
                                                                                                                                                                                                                                                                                                                /organism-"synthetic construct"
/mol_type-"genomic DNA"
/db_xref-"taxon:32630"
                                                                                                                                                                                                                                                                                                                                                                                                            /note="unnamed protein product"
                     ONA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           425
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Indels:
             2058 bp
Sequence 2 from Patent W00196875.
AX395309
AX395309.1 GI:21066308
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              US-09-965-594-1 (1-182) x AX395309 (1-2058)
                                                                                                                                                                                                                                                                             Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Б
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             580
                                                                                                                                                                                                                                                                                                                                                                                                                              /codon_start=1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 6.56e-68
946.00
100.008
98.908
99.278
                                                                                                                                                artificial sequences.
                                                                                                               synthetic construct
synthetic construct
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             634 C
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Percent Similarity:
Best Local Similarity:
Query Match:
DB:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           419
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Alignment Scores:
Pred. No.:
               LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
                                                                                                                                                                                                                                                                                                     source
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           BASE COUNT
ORIGIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ..
9
                                                                                                                                                                    REFERENCE
AUTHORS
                                                                                                                                                                                                                                        JOURNAL,
                                                                                                                                                                                                                                                                                                                                                                                             CDS
                                                                                                                                                                                                                                                                                 FEATURES
 4X395309
                                                                                                                                                                                                                          TITLE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Score:
```

Q

g

Ω

ò

q

ö

qq

ò

ò

q

g

õ

ò

g ò a

```
/codol_stait=1
/product=*NS3 protease"
/product=*NS3 protease"
/protein_id=*AAK54543.1*
/db_xxef=*C01:14150561"
/translation=*APITAYAQOTHGLLGC11TSLTGRDKNQVEGEVQIVSTAAQTFL
/translation=*APITAYTASPKGPYIQMYTNVDKDLVGWAPAGGSRSLTPCTCGSS
DLYLVTRHADVIPWRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVCT
RGVAKAVDFIPVENLETTMRS"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   VRL 03-JUL-2002
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Hepatitis C virus
Hepatitis C virus
Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                                         240
                                                                                                                                                                                                                                                                                                                                                                                                                                                   300
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              360
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   141
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        420
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            161
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 480
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       181
                                                                                                                                                                                                                                                                                                                                                                                                                               101
                                                                                                                                                                                                                                                                                                                                                                                                                                                                        121
                                                                                                                                                                                                                                                                                                          81
                                                                                                                                                                                                                                                                          9
                                                                                                                                                                                                                                                                                                41
                                                                                                                                                                                                                                                                                                                                          61
                                                                                                                                                                                                                                                                                                                                                                                                                             301 recressacertractrestaceassearscesarsteratacessescesses
                                                                                                                                                                                                                                                                                                                                                                                    62 ThrargThrIleAlaSerProLysGlyProValIleGlnMetTyrThrAsnValAspLys
                                                                                                                                                                                                                                                                                                                                                                                                 122 AspSerArgGlySerLeuLeuSerProArgProlleSerTyrLeuLysGlySerSerGly
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               GlnThrPheLeuAlaThrCysIleAsnGlyValCysTrpThrValTyrHisGlyAlaGly
                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SerSerAspLeuTyrLeuValThrArgHisAlaAspValIleProValArgArgArgGly
                                                                                                                                                                                                                                                                                                22 LeuThrGlyArgAspLysAsnGlnValGluGlyGluValGlnIleValSerThrAlaAla
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     AF369235 543 bp RNA linear VRL
Hepatitis C virus Pt.1Y NS3 protease gene, partial cds.
AF369235
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     543
1180
0
0
                                                                                                                                                    Length:
Matches:
Conservative:
Mismatches:
Indels:
                                                                                                           106
                                                                                                                                                                                                                               US-09-965-594-1 (1-182) x AF369218 (1-543)
<1. .>543
/note-"polyprotein"
                                                                                                          158 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  AF369235.1 GI:14150594
                                                                                                                                                   1.89e-68
945.00
100.00%
99.45%
99.16%
                                                                                                           166 c
                                                                                                                                                                        Percent Similarity:
Best Local Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Ser 182
                                                                                                            ø
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      543
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      541 TCC
                                                                                                                                           Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  182
                                                                                                                                                                                                                                                                                                                                                                                                                                                                         102
                                                                                                                                                                                                                                                                                                                                          42
                                                                                                                                                                                                Query Match:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             DEFINITION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ORGANISM
                                                                                                                                                    No.
                                                                                                           BASE COUNT
ORIGIN
                                                                                                                                           Alignment
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ACCESSION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           RESULT 4
AF369235
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  VERSION
KEYWORDS
SOURCE
  CDS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  LOCUS
                                                                                                                                                                 Score
                                                                                                                                                       Pred
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      a
                                                                                                                                                                                                                                                                                                                                                                                                        qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                 Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                        õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ДQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                  qq
                                                                                                                                                                                                                                                                                                                                          ö
                                                                                                                                                                                                                                                                                                                                                              a
                                                                                                                                                                                                                                                                                                                                                                                   οχ
                                                                                                                                                                                                                                                                                                                                                                                                                                5
                                                                                                                                                                                                                                                     õ
                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Genetic diversity and response to IFN of the NS3 protease gene from clinical strains of the hepatitis C virus Arch. Virol. 147 (7), 1385-1406 (2002)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1 (bases 1 to 543)
Holland-Staley, C.A., Kovari, L.C., Golenberg, E.M., Pobursky, K.J. and
                                                                                                                                                                                                                                                                                                                                                                                                                                                                         03-JUL-2002
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                   360
                                                                                                                                                                                                                                                                          140
                                                                                                                                                                                                                                                                                               420
                                                                                                                                                               240
                                                                                                                                                                                     100
                                                                                                                                                                                                         300
                                                                                                                                                                                                                               120
                                                                          120
                                                                                                                     180
          20
                                9
                                                     SerLeuThrGlyArgAspLysAsnGlnValGluGlyGluValG]nIleValSerThrAla 40
                                                                                                69
                                                                                                                                           80
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Direct Staley, C.A., Kovari, L.C., Golenberg, E. and Mayers, D.I.. Direct Submission Submission Infectious Disease Research, Henry Ford Health Systems, 2799 W. Grand Blvd. Rm 7045 E & R, Detroit, MI
                                                                                              101 GlySerSerAspLeuTyrLeuValThrArgHisAlaAspValIleProValArgArgArg
                                                                 MetAlaProIleThrAlaTyrAlaGlnGlnThrArgGlyLeuLeuGlyCysIleIleThr
                                                                                                                                                                                                                                                                                                                                                                                                                                                                         VRL
                                                                                                                                                                                                                                                                                                                                                                                                                                                           Hepatitis C virus Pt.4 NS3 protease gene, partial cds. AF369218
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   DNA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     00
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Viruges; SSRNA positive strand viruses, Hepacivirus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  /organism-"Hepatitis C virus"
/mol_type-"genomic RNA"
/isolate-"Pt.4"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   /db_xref="taxon:11103"
/note="type: 1A"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Hepatitis C virus
Hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                               546
                                                                                                                                                                                                                                                                                                                                                                                                           ArgSer 182
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Mayers, D.L.
                                                                                                                                                                                                                                                                                                                                                                                                                               AGGTCC
                                                                                                                                                                                                                                                                                                                                                                                                           181
                                                                                                                                                                                                                                                                                                                                                                                                                               541
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          source
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    DEFINITION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ORGANISM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      AUTHORS
TITLE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        REFERENCE
AUTHORS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ACCESSION
VERSION
                                                                                                                                                                                                                                                                                                                                                                                                                                                   RESULT 3
AF369218
LOCUS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   KEYWORDS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       MEDLINE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   PUBMED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             REFERENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             JOURNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            JOURNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               PEATURES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        TITLE
```

Db

ò

g

ò

qq δλ q õ g

ογ

g

ö

us-09-965-594-1.rge

```
Unclassified.
1 (bases 1 to 5360)
1 (chases 1 to 5360)
Citien, D.Y.
NANBV diagnostics and vaccines
Patent: US 6150087-A 53 21-NOV-2000;
Location/Qualifiers
1. 5360
                                                                                                                                    5360 bp
US 6150087.
                                                                                                                                                                                                                                                                                              1145
                                                                                                                                                                                                                                                                                                                                                                                         Gaps:
                                                                                                                                                                                                                                                                                                                                                                                                            US-09-965-594-1 (1-182) x AR118686 (1-5360)
                                                                                                                                                                                                                                                                                 /organism="unknown"
1623 c 1532 g
                                                                                                                                              Sequence 53 from patent
AR118686
AR118686.1 GI:14100596
                                                                                                                                                                                                                                                                                                                                  3.22e-67
943.00
100.00%
98.35%
98.95%
                                                                                                                                                                                                                                                                                                                                                                    Best Local Similarity:
Query Match:
                                                                                                                                                                                                                                                                                              1060 a
                                                                        Ser 182
                                                                                            541 TCC 543
                                                                                                                                     AR118686
                                                                                                                                                                                        Unknown
                                                                                                                                                                                                                                                                                                                                                           Percent Similarity:
                                                                                                                                                                                                                                                                                                                             Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1290
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1110
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1350
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               101
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         121
                                                                        182
                                                                                                                                    LOCUS
DEFINITION
                                                                                                                                                                                                                                                                           source
                                                                                                                                                                                                                                                                                           BASE COUNT
                                                                                                                                                                                                                                                                                                                                      Pred. No.:
                                                                                                                                                                                                   ORGANISM
                                                                                                                                                                                                                                AUTHORS
TITLE
JOURNAL
                                                                                                                                                         ACCESSION
VERSION
KEYWORDS
SOURCE
                                                                                                                                                                                                                        REFERENCE
                                                                                                                RESULT 5
AR118686
                                                                                                                                                                                                                                                                FEATURES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Q
                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               d
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     \delta
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       QC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                               ò
                                                   g
                                                                        ò
                                                                                            g
                                                                                                                                                                                                                                        /codo_start=1
/product="NS3 protease"
/product="NS3 protease"
/product="NS3 protease"
/db_xref="C1:111550.1"
/tb_xref="C1:11550.1"
/translation-"APITAYAQOTRGLLGCIITSLIGRDKNQVEGEVQIVSTAAQTFL
ATCINGVCWTVYHGAGTRIASPKGPVIQMYINVDKDLVGWPAPQGSRSLTPCTCGSS
DI;KLWTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVCT
RGVARAVDFIPVETTMRS"
109 c 159 g 103 t
                        Genetic diversity and response to IFN of the NS3 protease gene from clinical strains of the hepatitis C virus Arch. Virol. 147 (7), 1385-1406 (2002)
            and
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1.83
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               240
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      81
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                9
                                                                                (bases 1 to 543)
Holland-Staley, C.A., Kovari, L.C., Golenberg, E. and Mayers, D.L.
Direct Submission
Submitted (09-ARR-2001) Infectious Disease Research, Henry Ford
Health Systems, 2799 W. Grand Blvd. Rm 7045 E & R, Detroit, MI
                                                                                                                                                                                                                                                                                                                                                                                                                                                               1 (bases 1 to 543)
Holland-Staley,C.A., Kovari,L.C., Golenberg,E.M., Pobursky,K.J.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ThrargThrIlealaSerProLysGlyProValIleGlnMetTyrThrAsnValAspLys
                                                                                                                                                                                                                                                                                                                                                                     543
180
0
0
                                                                                                                                                                                                                                                                                                                                                                   Length:
Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                                                                 /organism="Hepatitis C virus"
/mol_type="genomic RNA"
/isolate="pt.ly"
                                                                                                                                                                                                                                                                                                                                                                                                                        Gaps:
                                                                                                                                                                                                 /db_xref="taxon:11103"
/note="type: 1A"
                                                                                                                                                                                                                                                                                                                                                                                                                                           (1-543)
                                                                                                                                              Location/Qualifiers
                                                                                                                                                                                                                                /note-"polyprotein"
                                                                                                                                                                                                                                                                                                                                                                                                                                           US-09-965-594-1 (1-182) x AF369235
                                                                                                                                                                                                                                                                                                                                                                  1,89e-68
945.00
100.00%
99.45%
99.16%
                                                                                                                                                                                                                        <1, .>543
                     Mayers, D.L.
                                                                                                                                                                                                                                                                                                                                                                                                  Best Local Similarity:
Query Match:
DB:
                                                                       12111414
                                                                                                                                                                                                                                                                                                                                                                                         Percent Similarity:
                                                                                                                                                                                                                                                                                                                                                           Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           62
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      122
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          361
                                                                                                                                                          source
                                                                                                                                                                                                                                                                                                                           BASE COUNT
ORIGIN
                                                                                                                                                                                                                                                                                                                                                                    Pred. No.:
                                                  JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS
                                                                                                      TITLE
JOURNAL
                                                                                                                                                                                                                        CDS
                                                                                                                                               FEATURES
                               TITLE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ď
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ογ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Op
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
```

```
GlyAspSerArgGlySerLeuLeuSerProArgProIlcSerTyrLeuLysGlySerSer 140
                                                                                                                     PAT 16-MAY-2001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           141 GlyGlyProLeuLeuCysProAlaGlyHisAlaValGlyIlePheArgAlaAlaValCys 160
930 CIGGCGCCCATCACGGCGIACGCCCAGCAGACAAGGGCCCICCIAGGGIGCAIAAICACC 989
                                                                                                                                                                                                                                                                                                                                                                                                                                  9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                   80
                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1 MetalaProlleThralaTyralaGlnGlnThrargGlyCeuLeuGlyCysIleIleThr
                                                                                                                                                                                                                                                                                                                                                                                                                                  41 AlaGlnThrPheLeuAlaThrCysIleAsnGlyValCysTrpThrValTyrHisGlyAla
                                                                                                                      linear
                                                                                                                                                                                                                                                                                   5360
179
0
0
                                                                                                                                                                                                                                                                                   Length:
Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                     DNA
```

Db	8	LOCUS 109328 5360 bp DNA linear PAT 02-DEC-1994 DEFINITION Sequence 8 from Patent WO 8904669. ACCESSION 109328	REFERENCE 1 (bases 1 to 5360) AUTHORS Houghton.M., Choo.QK. and Kuo.G. AUTHOR Patent: WO 8904669-A 8 01-JUN-1989;	e 1061 a	3.22e-67 Length: 943.00 Matches:	Percent Similarity: 100.00% Conservative: 3   Best Local Similarity: 98.35% Mismatches: 0   Query Match: 98.95% Mismatches: 0   DB: 6   Gaps: 0   Gaps: 0	Qy         1 MetalaprolleThralaTyralaGlnGlnThrArgGlyLeuLeuGlyCysIleIleThr 20           :::1	21 SerLeuthrGlyArgAspLysAsnGlnValGluGlyGluValGlnIleValSerthrAla 	Qy 41 AlaGinThrPheLeuAlaThrCysileAsnGlyValCysTrpThrValTyrHisGlyAla 60	0y 61 GlyThrArgThrIleAlaSerProLysGlyProValIleGinMetTyrThrAsnValAsp 80	Qy 81 LysaspLeuValGlyTrpProAlaProGlnGlySerArgSerLeuThrProCysThrCys 100 :::	Qy 101 GlySerSerAspLeuTyrLeuValThrArgHisAlaAspVal1leProValArgArgArg 120 HIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIII	0y 121 GlyaspserArgGlySerLeuLeuSerProArgProIleSerTyrLeuLysGlySerSer 140	Qy 141 GlyGlyProLeuLeuCysProAlaGlyHisAlaValGlyIlePheArgAlaAlaValCys 160	Oy 161 ThrargGlyValAlaLysAlaValAspPheIJeProValGluSerLeuGluThrThrMet 180	Oy 181 ArgSer 182
161 ThrArgGly           1410 ACCGTGGA	Oy 181 ArgScr 182   1/1/1/1   Db 1470 AGTCC 1475	RESULT 6 106434 LOCUS 106434 LOCUS 106434 Experiment Fig. 5360 bp DNA linear PAI 02-DEC-1994 ACCESSION 106434.1 GI:590311 VERSION 106434.1 GI:590311	_	AUTHORS Houghton M., Choo.OL. and Kuo.G.  TITLE Nanby diagnostics and vaccines JOURNAL Patent: EP 0318216 A1 48 31-MAY-1989; FEATURES Location/Qualifiers Source i. 5360	/organism-"unknown" BASE COUNT 1061 a 1623 c 1533 g 1143 t ORIGIN	Alignment Scores: 3.22e-67 Length: 5360 Pred. No.: 943.00 Matches: 179 Score: 100.00% Conservative: 3 Best Local Similarity: 98.35% Mismatches: 0	Indels: Gaps: (1-5360)	laProlleThrAla' 	Qy 21 SerLeuThrGlyArgASpLySASnGlnValGluGlyGluValGlnIleValSerThrAla 40 	<pre>Qy 41 AlaGlnThrPheLeuAlaThrCysIleAsnGlyValCysTrpThrValTyrHisGlyAla 60                                      </pre>	Qy 61 GlyThrargThrIleAlaSerProLysGlyProVallleGlnMetTyrThrasnValAsp 80 	Oy 81 LysaspLcuValGlyTrpProAlaProGlnGlySerargSerLeuThrProCysThrCys 100 :::	Oy 101 GlySerSerAspleuTyrLeuValThrArgHisAlaAspVall1eProValArgArgArg 120 	<pre>Qy 121 GlyAspSerArgGlySerLeuLeuSerProArgProlleSerTyrLeuLysGlySerSer 140</pre>	<pre>Qy 141 GlyGlyProLeuLeuCysProAlaGlyHisAlaValGly1lePhcArgAlaAlaValCys 160</pre>	Qy 161 ThrargGlyValAlaLySAlaValAspPheIleProVaiGluScrLeuGluThrHet 180 

us-09-965-594-1.rge

```
DNA
                                                                  1 (bases 1 to 6785)
Houghton, W. Choo, Q.-L. and Kuo, G.
Nanby diagnostics and vaccines
Patent: EP 0318216-A1 54 31-MAY-1989;
6785 bp
EP 0318216.
                                                                                                                                                                                                                      US-09-965-594-1 (1-182) x 106440 (1-6785)
                                                                                                     Location/Qualifiers
1. .6785
                                                                                                                   /organism-"unknown"
       Sequence 54 from Patent
106440
                                                                                                                                                            4.15e-67
943.00
100.00%
98.35%
                          GI:590312
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1743 AGGTCC 1748
                                                                                                                                                                     Score:
Percent Similarity:
1
Pest Local Similarity:
9
Ouery Match:
6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            181 ArgSer 182
                                                           Unclassified
                                                                                                                            1392 a
                         106440.1
                                          Unknown
                                                    Unknown
                                                                                                                                                      Alignment Scores:
Pred. No.:
                                                                                                                                                                                                                                                                                                                                         61
                                                                                                              source
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
                                                                           AUTHORS
TITLE
JOURNAL
                                                                                                                             BASE COUNT
                                                  ORGANISM
                                                                   REFERENCE
                                                                                                     FEATURES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              RESULT
                                                                                                                                     ORIGIN
                                                                                                                                                                                                                                      οy
                                                                                                                                                                                                                                                      q
                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                      a
                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                      q
                                                                                                                                                                                                                                                                                                                                                                                                       \delta
                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                       qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Ö,
                                                                                                                                                                                                                                                             1562
                                                                                                                                                                                                                                                                                                                                                                                                                                                                1623 GGGGGTCCGCTGTTGTGCCCCGCGGGCACGCCGTGGGCATATTTAGGGCCGCGGGGTGTGC 1682
                         PAT 16-MAY-2001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  141 GlyGlyProLeuLeuCysProAlaGlyHisAlaValGLyIlePheArgAlaAlaValCys 160
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   180
                                                                                                                                                                                                                                                                                                                                                                                                                                 101 GlySerSerAspLeuTyrLeuValThrArgHisAlaAspValIleProValArgArgArg 120
                                                                                                                                                                                                                                                                                                                                41 AlaGlnThrPheLeuAlaThrCysIleAsnGlyValCysTrpThrValTyrHisGlyAla 50
                                                                                                                                                                                                                                                                                                                                                                                                                                          ThrArgGlyValAlaLysAlaValAspPheIleProValGluSerLeuGluThrThrMet
                          linear
                                                                                                                                                                                                      m 000
                                                                                                                                                                                     Length:
Matches:
Conservative:
Mismatches:
Indels:
                          DN.
                                                                                                                                                    نه
                                                                                         Chien, D.Y.

Chien, D.Y.

NANBV diagnostics and vaccines
Patent: US 6150087-A 65 21-NOV-2000;
Location/Qualifiers
1. 6785
                         6785 bp 1 or 6150087.
                                                                                                                                                     1429
                                                                                                                                                                                                                              Gaps:
                                                                                                                                                                                                                                              US-09-965-594-1 (1-182) x AR118692 (1-6785)
                                                                                                                                         /organism-"unknown"
2050 c 1914 q
                                 patent
                         AR118692
Sequence 65 from patent
AR118692
AR118692.1 GI:14100562
                                                                                                                                                                                    4.15e-67
943.00
100.00%
98.35%
98.95%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                AGGTCC 1748
                                                                                   Unclassified.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ArgSer 182
                                                                                                                                                                                                     Percent Similarity:
Best Local Similarity:
                                                                                                                                                      ď
                                                                   Unknown.
                                                                            Unknown
                                                                                                                                                                               Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                 1323 (
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1743
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   161
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    181
                                                                                                                                                                                                                      Query Match:
DB:
                         LOCUS
DEFINITION
                                                                                                                                      source
                                                                   SOURCE
ORGANISM
                                                                                                    AUTHORS
TITLE
JOURNAL
                                                                                                                                                     BASE COUNT
                                                                                                                                                                                        S
S
                                          ACCESSION
VERSION
KEYWORDS
                                                                                            REFERENCE
        RESULT 8
AR118692
                                                                                                                              FEATURES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       RESULT
                                                                                                                                                             ORIGIN
                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                            q
                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                           οχ
                                                                                                                                                                                                                                                                                                                                            đ
                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                            a
                                                                                                                                                                                                                                                                                                                                                                                             δ
                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    a
```

```
1382
                                                                                                                                                                                                                                                                                                                                                          1442
                                                                                                                                                                                                                                                                                                                                                                                          1502
                                                                                                                                                                                                                                                                                                                                                                                                                          1562
                                                                                                                                                                                                                                                                                                                                                                                                                                          PAT 02-DEC-1994
                                                                                                                                                                                                                                                                                                                                                                           100
                                                                                                                                                                                                                                                                                                                                                                                                           120
02-DEC-1994
                                                                                                                                                                                                                                                                                                         09
                                                                                                                                                                                                                                                                                                                                          80
                                                                                                                                                                                                                                                                                                                                                  1323 GCCCAAACCTTCCTGGCAACGTGCATCAATGGGGTGTGCTGGACTGTACCACCACGGGGCC
                                                                                                                                                                                                                                                                                                         41 AlaGlnThrPheLeuAlaThrCysIleAsnGlyValCysTrpThrValTyrHisGlyAla
                                                                                                                                                                                                                                                                                                                                          GlyThrArgThrIleAlaSerProLysGlyProVallleGlnMetTyrThrAsnValAsp
                                                                                                                                                                                                                                                                                                                                                                                                          101 GlySerSerAspLeuTyrLeuValThrArgHisAlaAspVallleProValArgArgArg
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   linear
linear
                                                                                                                                                             6785
179
3
0
0
                                                                                                                                                             Length:
Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  LOCUS 109329 6785 bp
DEFINITION Sequence 10 from Patent WO 8904669
```

```
Unclassified.
1 (bases 1 to 7310)
                                                                                                                                                                                                                                                                                                                                                                                                                            181 ArgSer 182
       Unclassified.
                                                                                                         Percent Similarity:
Best Local Similarity:
                                                                1495 a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Unknown.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Unknown
                                                                                     Alignment Scores:
                                                                                                                                                                                                                                                               1908
                                                                                                                                                                                                                                                                                                                       2028
                                                                                                                                                                                                                                                                                                                                                    2088
                                                                                                                                                                                                                                                                                                                                                                                2148
                                                                                                                                                                                                                                                                                                         101
                                                                                                                                                                                                                                                                                                                                      121
                                                                                                                       Query Match:
DB:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                            DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ORGANISM
                                                                                             Pred. No.:
                                                                                                                                                                                                                                                                                                                                                                                                                                                        RESULT 12
I09331
 ORGANISM
              REFERENCE
AUTHORS
TITLE
JOURNAL
                                                                BASE COUNT
ORIGIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                REFERENCE
                                           FEATURES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1,000
                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                               qq
                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                            ò
                                                                                                                                                                        qq
                                                                                                                                                                                        ò
                                                                                                                                                                                                   q
                                                                                                                                                                                                                 õ
                                                                                                                                                                                                                                 a
                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                            a
                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                       q
                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                      οg
                                                                                                                                                                                                                                                                                                                                     03
                                                                                                                                                                                                                                                                                                                                                                  ó
                                                                                                                                                                                                                                                                                                                                                                                                                   1443 CAGACCTTGTGGGCTGGCCGCTCGCCCAAGGTAGCCGCTCATTGACACCCTGCACTTGC 1502
                                                                                                                                                                                                                                                                                                                             PAT 16-MAY-2001
                                                                                                                                                                                               1262
                                                                                                                                                                                                             81 LysAspLeuValGlyTrpProAlaProGlnGlySerArgSerLeuThrProCysThrCys 100
                                                                                                                                                                                                                                                                                                                                                                                       GlyGlyProLeuLeuCysProAlaGlyHisAlaValGlyIlePheArgAlaAlaValCys 160
                                                                                                                                                                                                                                                                      80
                                                                                                                                                                                 GlyThrArgThrIleAlaSerProLysGlyProValIleGlnMeLTyrThrAsnValAsp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            linear
                                                                                                                 6785
179
3
0
0
                                                                                                                Length:
Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             DNA
                                  Unclassified.
1 (bases I to 6785)
Houghton,M., Choo,Q.-K. and Kuo,G.
Patent: WO 8904669-A 10 01-JUN-1989;
Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Sequence 74 from patent US 6150087.
AR118696
AR118696.1 GI:14100606
                                                                                     1429
                                                                                                                                                      Gaps:
                                                                                                                                                                   US-09-965-594-1 (1-182) x I09329 (1-6785)
                                                                           /organism-"unknown"
2050 c 1914 a
                                                                                                                 4.15c-67
943.00
100.00%
98.35%
98.95%
        GI:587964
                                                                                                                                                                                                                                                                                                                                                                                                                                                          111111
AGGTCC 1748
                                                                                                                                                                                                                                                                                                                                                                                                                                                 181 ArgSer 182
                                                                                                                                Percent Similarity:
Best Local Similarity:
                                                                                      ۵
109329
109329.1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Unknown
                      Unknown
                              Unknown
                                                                                                            Scores:
                                                                                                                                                                                                                                                                                                                                                                         1563
                                                                                                                                                                                                                                                                                                                                                                                                                                                              1743
                                                                                                                                                                                                                                                                      61
                                                                                                                                                                                                                                                                                                                                                            121
                                                                                                                                                                                                                                                                                                                                                                                       141
                                                                                                                                                Query Match:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   DEFINITION
                                                                         source
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ACCESSION
VERSION
KEYWORDS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                             RESULT 11
AR118696
LOCUS
                              ORGANISM
                                                   AUTHORS
JOURNAL
FEATURES
                                                                                      BASE COUNT
                                                                                                            Alignment (
Pred. No.:
 ACCESSION
                                            REFERENCE
        VERSION
KEYWORDS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                SOURCE
                       SOURCE
                                                                                             DRIGIN
                                                                                                                          Score:
                                                                                                                                                                                                                                                                                                                                                                       Q
                                                                                                                                                                                                                                                                                                                                                                                   õ
                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                        qq
                                                                                                                                                                                                                                                                  0 PP
                                                                                                                                                                                                                                                                                                            g S
                                                                                                                                                                                                                                                                                                                                         a
                                                                                                                                                                                                                                                                                                                                                       ô
                                                                                                                                                                                               q
                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                           oy
Dp
                                                                                                                                                                                  ò
```

```
2147
                                                                                                                                                                                                                              1847
                                                                                                                                                                                                                                                                                                   1961
                                                                                                                                                                                                                                                                                                                                     2027
                                                                                                                                                                                                                                                                                                                                                                       2087
                                                                                                                                                                                                                                                                                                                                                                                         140
                                                                                                                                                                                                                                                                                                                                                                                                                                                             PAT 02-DEC-1994
                                                                                                                                                                                                                                                                  1907
                                                                                                                                                                                                                                                                                                                                                       120
                                                                                                                                                                                                                                                                                                                                                                                                                             141 GlyGlyProLeuLeuCysProAlaGlyHisAlaValGlyIlePheArgAlaAlaValCys 160
                                                                                                                                                                                                                                                                                                                     100
                                                                                                                                                                                                                                                09
                                                                                                                                                                                                                                                                                    80
                                                                                                                                                                             20
                                                                                                                                                                                                              40
                                                                                                                                                                                                                                                                                                                                                      GlyAspSerArgGlySerLeuLeuSerProArgProIleSerTyrLeuLysGlySerSer
                                                                                                                                                                                                                                                                                                                                                                                                                                     61 GlyThrArgThrIleAlaSerProLysGlyProValIleGlnMetTyrThrAsnValAsp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       linear
                                                                                               Length:
Matches:
Conservative:
Mismatches:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       DNA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   7310 bp
Sequence 15 from Patent WO 8904669.
109331.1 GI:587966
                                                                                                                                 Indels:
Gaps:
                                                                                                                                                           US-09-965-594-1 (1-182) x AR118696 (1-7310)
1 (bases 1 to 7310)
(Chieh, Dr. Y.
MRNBV diagnostics and vaccines patent: US 6150087-A 74 21-NOV-
                                                  /organism="unknown"
2220 c 2056 g
                                   Location/Qualifiers
                                                                                              4.5e-67
943.00
100.008
98.35%
98.95%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   2268 AGGTCC 2273
```

æ

```
/protein_id-"AAA45677.1"
/db_xref**G1:329876*
/tanslaton-*GCPEPTSPWTVGTTDFDQGWGPISYANGSGPDQRPYCWHYPPRPC
/translation-*GGPEPTSPWTVGTTDRSGAPTYSWGENDTDVFVLNNTRPPLGNWFGC
TWANSTGFTKVCGAPFCVIGGAGNNTLHCPTDCFRKHPDATYSRCGSGPWITPRCLVD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          QEMGGNITRVESENKYVILDSFDPLVAEEDEREISVPAEILERSRRFAQALPVMARPD
YNPPLVETWKKPDYEPPVVHGCPLPPPKSPPVPPRKKRTVVLTESTLSTALAELATR
SFGSSSTSGITGDNTTISSEPAPSGCPPDSDAESYSSMPPLEGEPGDPDLSDGSWSTV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 PDLGVRVCEKMALYDVVTKLPLAVMGSSYGFQYSPGQRVEFLVQAWKSKKTPMGFSYD
TRCFDSTVTESDIRTEEAIYQCCDLDPQARVAIKSLTERLYVGGPLTNSRGENGGYRR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     RAFTEAMTRYSAPPGDPPOPEYDLELITSCSSNVSVAHDGAGKRVYYLTRDPTTPLAR
AAWETARHTPVNSWLGNIIMFAPTLWARMILMTHFFSVLIARDQLEQALDCEIYGACY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         QRQKKVTFDRLQVLDSHYQDVLKEVXAAASKVKANLLSVEEACSLTPPHSAKSKFGYG
AKDVRCHARKAVTHINSVWKDLLEDNVTPIDTTIMAKNEVFCVQPEKGGRKPARLIVF
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CRASGVLITSCGNTLICYIKARAACRAAGLQDCIMLVCGDDLVVICESAGVQEDAASL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SSEANAEDVVCCSMSYSWTGALVTPCAAEEQKLPINALSNSLLRHHNLVYSTTSRSAC
                                                                                                                                                                                                                                           YPYRLWHYPCTINYTIFKIRMYVGGVEHRLEAACNWTRGERCDLEDRDRSELSPLLLT
     Original source text: Hepatitis C virus, cDNA to viral RNA, clones K9-1 through 15e, isolated from chimpanzee (individual 910) blood
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1 MetAlaProlleThrAlaTyrAlaGlnGlnThrArgGlyLeuLeuGlyCys1leIleThr 20
                            plasma.
Draft entry and printed sequence for [1] kindly submitted by
M.Houghton, 22-FEB-1990. Chiron Corporation, 4560 Horton Street,
Emeryville CA 94608
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           7310
179
3
0
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Matches:
Conservative:
Mismatches:
Indels:
                                                                                          1. .7310
/organism="Hepatitis C ./mol_type="genomic RNA"
/db_xref="taxon:11103"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 US-09-965-594-1 (1-182) x HPCPOLYP (1-7310)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         2058 9
                                                                             Location/Qualifiers
                                                                                                                                            <1. .>7310
/note="polyprotein"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            SIEPLDLPPIIQRL'
                                                                                                                                                                     /codon_start=3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          4.5e-67
943.00
100.00%
98.35%
98.95%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         2218 c
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Best Local Similarity:
Query Match:
DB:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1495
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Percent Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Alignment Scores:
                                                                                             source
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Pred. No.:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          BASE COUNT
                                                                                                                                           CDS
                                                                              FEATURES
     COMMENT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ORIGIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Score:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ð
                                                                                                                                                                                                                         2027
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    2087
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         2207
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     02-AJG-1993
                                                                                                                                                                                                                                                                                                                                                                                                                             100
                                                                                                                                                                                                                                                                                                                                                                                                                                                                       160
                                                                                                                                                                                                                                                                                                                                                                                                                         141 GlyGlyProLeuLeuCysProAlaGlyHisAlaVaiGlyflcPheArgAlaAlaValCys
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        2148 GGGGGTCCGCTGTTGTTGTGCCCCGCGGGGCACGCCGTGGGCATATTTAGGCCGCGCGGGGTGTGC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     VRL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1 (bases 1 to 7310)
Choo,O.-L., Richman,K. and Han,J.
The nucleotide sequence of the Hepatitis C viral genome
Unpublished (1990)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Hepatitis C virus polyprotein gene, partial cds. M32084.1 G1:329875
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          no DNA
                                                                                                                 Length:
Matches:
Conservative:
Mismatches:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Hepatitis C virus
Hepatitis C virus
Viruses: ssRNA positive-strand viruses,
Houghton,M., Choo,Q.-K. and Kuo,G. Patent: WO 8904669-A 15 01-JUN-1989; Location/Qualifiers
                                                                 1539
                                                                                                                                                                   Indels:
                                                                                                                                                                                                    US-09-965-594-1 (1-182) x 109331 (1-7310)
                                                    /organism="unknown"
2218 c 2058 g
                                                                                                                 4.5c-67
943.0c
100.00%
98.35%
98.95%
                                          .7310
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          AGGTCC 2273
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Hepacivirus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           181 Argser 182
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   polyprotein.
                                                                                                                          Score:
Percent Similarity:
Best Local Similarity:
                                                                   q
                                                                                                     Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 2268
                                                                                                                                                               Query Match:
DB:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    HPCPOLYP
LOCUS
DEFINITION
ACCESSION
VERSION
                                          source
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  KEYWORDS
SOURCE
ORGANISM
                                                                BASE COUNT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               REFERENCE
AUTHORS
                                                                                                                 N.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        TITLE
JOURNAL
              JOURNAL
 AUTHORS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           RESULT 13
                            FEATURES
                                                                             ORIGIN
                                                                                                                 Pred.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                QQ
                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                 a
                                                                                                                                                                                                                                                                                                                                              Q
                                                                                                                                                                                                                                                                                                                                                                                                q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             음
                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                  .8
                                                                                                                                                                                                                                                                                                                                                                                                                                              qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ç
```

2973

80

100

2913

09

3093

140

3153

160

3213

180

```
PAT 16-MAY-2001
                                                                                                                   161 ThrArgGlyValAlaLysAlaValAspPheIleProValGluSerLeuGluThrThrMet
                            GlyGlyProLeuLeuCysProAlaGlyHisAlaValGlyIlePheArgAlaAlaValCys
                                                                                        LysAspLeuValGlyTrpProAlaProGlnGlySerArgSerLeuThrProCysThrCys
                                                          GlyThrArgThrIleAlaSerProLysGlyProValIleGlnMetTyrThrAsnValAsp
                                                                                                                                                                                                                                                                                     linear
                                                                                                                                                                                                                                                                                                                                                                                                                                8987
179
3
0
0
                                                                                                                                                                                                                                                                                                                                                                                                                                        Matches:
Conservative:
Mismatches:
                                                                                                                                                                                                                                                                                     DNA
                                                                                                                                                                                                                                                                                                                                            1 (bases 1 to 8987)
Chien, D.Y.
NANBV diagnostics and vaccines
Patent: US 6150087-A 137 21-NOV-2000;
Location/Qualifiers
                                                                                                                                                                                                                                                                                      7 bp DN
6150087.
                                                                                                                                                                                                                                                                                                                                                                                                                                 Length:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    US-09-965-594-1 (1-182) x AR118728 (1-8987)
                                                                                                                                                                                                                                                                                                                                                                                             /organism="unknown"
2735 c 2547 g
                                                                                                                                                                                                                                                                                           Sequence 137 from patent US
AR118728
AR118728.1 GI:14100638
                                                                                                                                                                                                                                                                                                                                                                                                                                5.62e-67
943.00
100.00%
98.35%
98.95%
                                                                                                                                                                                                                                                  3274 AGGTCC 3279
                                                                                                                                                                                                                                                                                                                                         Unclassified.
                                                                                                                                                                                                                                         181 ArgSer 182
                                                                                                                                                                                                                                                                                                                                                                                                                                                Percent Similarity:
Best Local Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                     Ø
                                                                                                                                                                                                                                                                                     AR118728
                                                                                                                                                                                                                                                                                                                                                                                                    1807
                                                                                                                                                                                                                                                                                                                          Unknown
                                                                                                                                                                                                                                                                                                                                  Unknown
                                                                                                                                                                                                                                                                                                                                                                                                                           Alignment Scores:
Pred. No.:
                                                                                                                                                                                141
                                                                                       81
                                                          61
                                                                                                                                                                                                                                                                                            DEFINITION
ACCESSION
VERSION
                                                                                                                                                                                                                                                                                                                                                      AUTHORS
TITLE
JOURNAL
FEATURES
                                                                                                                                                                                                                                                                                                                                  ORGANISM
                                                                                                                                                                                                                                                                                                                                                                                                     BASE COUNT
                                                                                                                                                                                                                                                                       RESULT 15
AR118728
                                                                                                                                                                                                                                                                                                                                                 REFERENCE
                                                                                                                                                                                                                                                                                                                  KEYWORDS
                                                                                                                                                                                                                                                                                                                                                                                                            ORIGIN
                                                                                                                                                                                                                                                                                                                                                                                                                                          Score:
                                                                                                                                                                                                                                                                                      LOCUS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              a
                                                                                                                                                                                                                                                       qq
           q
                         ò
                                          qq
                                                          ò
                                                                     q
                                                                                       ò
                                                                                                    q
                                                                                                                  ô
                                                                                                                                 qu
                                                                                                                                                 ò
                                                                                                                                                              qq
                                                                                                                                                                              ^{\circ}
                                                                                                                                                                                            Db
                                                                                                                                                                                                           ò
                                                                                                                                                                                                                         ΩÇ
                                                                                                                                                                                                                                         à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          2027
                                                                                                                                                         2087
                                                                                                                                                                                      2:47
                                                                                                                                                                                                                    2207
                                                                                                                                                                                                                                 PAT 16-MAY-2001
                                                                                                                                                                        140
                                                                                                                                                                                                     160
                                                                                                              100
                                                                                                                                           GlySerScrAspLcuTyrLcuValThrArgHisAlaAspValIleProValArgArgArg 120
       CTGGCGCCCATCACGCCGTACGCCCAGCACAAGGGGCCTCCTAGGGTGCATAATCACC 178
                       40
                                                                                                              GlyGlyProLeuLeuCysProAlaGlyHisAlaValGlyIlePheArgAlaAlaValCys
                       SerLeuThrGlyArgAspLysAsnGlnValGluGlyGluValGlnIleValSerThrAla
linear
                                                                                                                                                                                                                                                                                                                                                                                                                                                        8316
179
3
0
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                      Length:
Matches:
Conservative:
Mismatches:
Indels:
Gaps:
                                                                                                                                                                                                                                                                                                             DNA
                                                                                                                                                                                                                                                                                                                                                                                                                           ĻĴ
                                                                                                                                                                                                                                                                                                                                                                     1 (bases 1 to 8316)
Chien,D.Y.
NANBV diagnostics and vaccines
Patent: US 6150087-A 88 21-NOV-2000;
                                                                                                                                                                                                                                                                                                        Sequence 88 from patent US 6150087.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           JS-09-965-594-1 (1-182) x AR118703 (1-8316)
                                                                                                                                                                                                                                                                                                                                                                                                     Location/Qualifiers
1. .8316
                                                                                                                                                                                                                                                                                                                                                                                                                   /organism="unknown"
                                                                                                                                                                                                                                                                                                                                                                                                                          2345 9
                                                                                                                                                                                                                                                                                                                                  AR118703.1 GI:14100613
                                                                                                                                                                                                                                                                                                                                                                                                                                                      5.17e-67
943.00
100.00%
98.35%
98.95%
                                                                                                                                                                                                                                                                                                                                                                                                                          2529 c
                                                                                                                                                                                                                                                                         Unclassified.
                                                                                                                                                                                                                                                                181 ArgSer 182
                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Percent Similarity:
Best Local Similarity:
Query Match:
                                                                                                                                                                                                                                                                                                                                                                                                                           a
                                                                                                                                                                                                                                                                                                                                                 Unknown.
                                                                                                                                                                                                                                                                                                                                                                                                                           1671
                                                                                                                                                                                                                                                                                                                                                          Unknown
                                                                                                                                                                                                                                                                                                                                                                                                                                                 Alignment Scores:
                                                                                                                                                                                                      141
                                                                                                                                                                                                                                                                               2268
                       21
                                                                                                                                            101
                                                                                                                                                                                                                                                                                            RESULT 14
AR118703
LOCUS
DEFINITION
                                                                                                                                                                                                                                                                                                                                                                                                             source
                                                                                                                                                                                                                                                                                                                                                 SOURCE
ORGANISM
                                                                                                                                                                                                                                                                                                                                                                                                                                                         ..
9
                                                                                                                                                                                                                                                                                                                          ACCESSION
VERSION
KEYWORDS
                                                                                                                                                                                                                                                                                                                                                                                                                           BASE COUNT
ORIGIN
                                                                                                                                                                                                                                                                                                                                                                       REFERENCE
                                                                                                                                                                                                                                                                                                                                                                                              JOURNAL
                                                                                                                                                                                                                                                                                                                                                                               AUTHORS
                                                                                                                                                                                                                                                                                                                                                                                                     FEATURES
                                                                                                                                                                                                                                                                                                                                                                                     TITLE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                Score:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                   Q
                                                              Q G
                                                                                             q
                                                                                                          o d
                                                                                                                                                        QQ
                                                                                                                                                                    ò
                                                                                                                                                                                    Op
                                                                                                                                                                                                    οy
                                                                                                                                                                                                                 qq
                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                              δ
                                                                                                                                                                                                                                                                            qq
                      ò
                                                 ò
                                                                                                                                          ò
```

Search completed: August 31, 2003, 00:45:16 Job time: 2382.6 secs

Sequence encoded i Peptide encoded by Protein encoded by HCV-1 polyprotein. Amino acid sequenc Polyprotein encode Compiled HCV seque

Hepatitis C virus Hepatitis C virus

Composite hepatiti

HCV polyprotein.
HCV polyprotein.
HCV polyprotein.
HCV polyprotein la

Hepatitis C virus Hepatitis C virus Hepatitis C virus Hepatitis C virus Hepatitis C virus

Hepatitis C

virus

Hepatitis C Hepatitis C

```
Modified hepatitis C virus (HCV) NS3 protease comprising at least 1 substitution of a hydrophobic alpha-helix 0 amino acid to a hydrophilic
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Hepatitis; NS3 protease; viral replication; chronic liver disease; liver failure; liver cancer.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Goldfarb V;
                                                                                                                                                                                                                                                                                                                                                                          ALIGNMENTS
                                                                                                                                                                                                       AAE21841
AAE21842
AAE21843
                                                                                                                                                                                                                                                                                  AAE19908
AAE19919
AAE19920
                            AAB18540
AAR34009
                                                       AAX14975
AAB18541
AAR21519
                                                                                                                 AAW40038
AAE22049
                                                                                                                                    AAU84597
AAE21847
                                                                                                                                                                                   AAE21839
AAE21840
                                                                                                                                                                                                                                   AAE21844
AAE21845
                                                                                                                                                                                                                                                                                                                         AAE19922
AAE19923
AAE19924
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Zhang Y,
                                                                                                                                                       AAE19905
                                                                                                                                                                                                                                                       AAE21846
                                                                                                                                                                                                                                                               AAE19900
                                                                                                                                                                                                                                                                                                                 AAE19921
                                                                                                                                                                                                                                                                                                                                                                                                                        AAB15211 standard; protein; 182 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               BRIM ) BRISTOL-MYERS SQUIBB CO.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Hepatitis C virus NS3 protease.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Wittekind M, Weinheimer S,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             99US-0115271.
 06-JAN-2000; 2000WO-US00345
                                                                                                                                                                                                                                                                                                                                                                                                                                                              19-DEC-2000 (first entry)
WPI; 2000-465976/40.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Hepatitis C virus
 N-PSDB; AAA70344
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    WO200040707-A1.
                                     08-JAN-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      13-JUL-2000
AAB15211;
                                                                                                                                                                                                                                                                                                                                                                                                       RESULT 1
                                                                                                                                                                                                                                                                                                                                                                                                                 AAB1521
 Hepatitis C virus
HCV-1 NS3/4a confo
HCV genomic amino
Hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Sequence encoded i
Protein sequence o
Peptide encoded by
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          HCV-1 NS3/4a mutan
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Hepatitis C virus
                                                             August 30, 2003, 17:42:58 ; Search time 41.2251 Seconds (without alignments) 700.745 Million cell updates/sec
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.
                                                                                                                                                                                                                                                                            A_Geneseq_19Jun03:*

| SIDSI/gcgdata/qeneseq/geneseqp-embl/AA1980.DAT:*
| SIDSI/qcgdata/qeneseq/geneseqp-embl/AA1981.DAT:*
| SIDSI/qcgdata/qeneseq/geneseqp-embl/AA1982.DAT:*
| SIDSI/qcgdata/qeneseq/geneseqp-embl/AA1983.DAT:*
| SIDSI/qcgdata/qeneseq/geneseqp-embl/AA1984.DAT:*
| SIDSI/qcgdata/qeneseqy-embl/AA1985.DAT:*
| SIDSI/qcgdata/qeneseqy-embl/AA1985.DAT:*
| SIDSI/qcgdata/qeneseqy-embl/AA1985.DAT:*
| SIDSI/qcgdata/qeneseqy-embl/AA1989.DAT:*
| SIDSI/qcgdata/qeneseqy-embl/AA1989.DAT:*
| SIDSI/qcgdata/qeneseqy-embl/AA1989.DAT:*
| SIDSI/qcgdata/qeneseqy-embl/AA1983.DAT:*
| SIDSI/qcgdata/qeneseqy-embl/AA1983.DAT:*
| SIDSI/qcgdata/qeneseqy-embl/AA1983.DAT:*
| SIDSI/qcgdata/qeneseqy-embl/AA1993.DAT:*
| SIDSI/qcgdata/qeneseqy-embl/AA1993.DAT:*
| SIDSI/qcgdata/qeneseqy-embl/AA1993.DAT:*
| SIDSI/qcgdata/qeneseqy-embl/AA1993.DAT:*
| SIDSI/qcgdata/qeneseqy-embl/AA1999.DAT:*
| SIDSI/qcgdata/qeneseqy-embl/AA1999.DAT:*
| SIDSI/qcgdata/qeneseqy-embl/AA1999.DAT:*
| SIDSI/qcqdata/qeneseqy-embl/AA1999.DAT:*
| SIDSI/qcqdata/qeneseqy-embl/AA2001.DAT:*
| SIDSI/qcqdata/qeneseqy-embl/AA2001.DAT:*
| SIDSI/qcqdata/qeneseqy-embl/AA2001.DAT:*
| SIDSI/qcqdata/qeneseqy-embl/AA2003.DAT:*
| SIDSI/qcqdata/qeneseqy-embl/AA2003.DAT:*
| SIDSI/qcqdata/qeneseqy-embl/AA2003.DAT:*
| SIDSI/qcqdata/qeneseqy-qeneseqp-embl/AA2003.DAT:*
| SIDSI/qcqdata/qeneseqy-qeneseqp-embl/AA2003.DAT:*
| SIDSI/qcqdata/qeneseqy-qeneseqp-embl/AA2003.DAT:*
| SIDSI/qcqdata/qeneseqy-qeneseqp-embl/AA2003.DAT:*
| SIDSI/qcqdata/qeneseqy-qeneseqp-embl/AA2003.DAT:*
                                                                                                                         1 MAPITAYAQQTRGLLGCIIT......GVAKAVDFIPVESLETTMRS 182
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Description
                                                                                                                                                                                             1107863
        GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.
                                                                                                                                                                                            Fotal number of hits satisfying chosen parameters:
                                                                                                                                                                          1107863 seqs, 158726573 residues
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                SUMMARIES
                                                                                                                                                                                                                                           Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries
                                               - protein search, using sw model
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         AAE18689
AAU76377
ABG72261
AAR40120
AAR51170
AAP92041
AAP90158
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               AAB15211
                                                                                                                                            BLOSUM62
Gapop 10.0 , Gapext 0.5
                                                                                                                                                                                                                length: 0
length: 2000000000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             DB
                                                                                                     US-09-965-594-1
953
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Ouery
Match Length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  686
686
3011
609
1766
1786
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              100.0
99.3
99.3
99.1
99.0
99.0
                                                                                                                                                                                                               DB seq 1
                                                                                                       Title:
Perfect score:
                                                                                                                                             Scoring table:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Score
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              OM protein
                                                                                                                           sedneuce:
                                                                                                                                                                          Searched:
                                                                                                                                                                                                                Minimum |
Maximum |
                                                                                                                                                                                                                                                                                   Database
                                                                 Run on:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             4500000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Result
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Š
```

a

```
Hepatitis C virus Synthetic.
                                   N-PSDB; AAD29795
                                                                                                                                                                                                                                                                                                                                                                                                                                                               181 RS 182
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      RS 182
Medina-Selby A;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     WO200196870-A2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     08-MAY-2002
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             20-DEC-2001
                                                                                                                                                                                                                                         Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            RESULT 3
qq
                                                                                                                                                                                                                                                                                                                                                                                     a
                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                               à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                ô
                                                                                                                                                                                                                                                                                                                                                                                                                ð
                                                                                                                                                                                                                                                  ó
                                                        The present sequence is the Hepatitis C virus (HCV) NS3 protease enzyme. This protein is essential for the replication of the virus, acting to cleave its replicative proteins from the polyprotein produced from the HCV genome. NS4A is also needed for this process and inhibitors of the HCV genome. NS4A is also needed for this process and inhibitors of the useful as HCV can lead to chronic liver disease such as cirrhosis. Invertailure and liver cancer. The present invention concerns a cumber of failure and NS3 matarits and NS3 NS4A faison proteins which can be used to identify inhibitors of this type, as well as enabling structural studies of the protease and protease inhibitor complexes.
                                                                                                                                                                                                                                                                                                                          GTRTIASPKGPVIQMYTNVDKDLVGWPAPQGSRSIJTPCTCGSSDLYLVTRHADVIPVRRR 120
                                                                                                                                                                                                                                                                                                                                       121 GDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVCTRGVAKAVDFIPVESLETTM 180
                                                                                                                                                                                                                                                                                                                                                                                       121 GDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVCTRCVAKAVDFIPVESLFTTM 180
                                                                                                                                                                                                                                                                                        1 MAPITAYAQQTRGLLGCIITSLIGRDKNQVEGEVQIVSTAAQTFLATCINGVCWTVYHGA 60
                                                                                                                                                                                                                                                                          1 MAPITAYAQQTRGLLGCIITSLTGRDKNQVEGEVQIVSTAAQTFLATCINGVCWTVYHGA 60
amino acid, useful for screening inhibitors that may treat hepatitis C
                                                                                                                                                                                                                                                     Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ä
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Hepatitis C virus; NS3/4a antigen; HCV infection; mutant; mutein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Coit
                                                                                                                                                                                                                                                  0;
                                                                                                                                                                                                                         Length 182;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Tandeske L, George-Nasciemento C,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     /note= "Wild type Thr substituted with Pro"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             "Wild type Ser substituted with Ile"
                                                                                                                                                                                                                                                  Indels
                                                                                                                                                                                                                         Score 953; DB 21;
Pred. No. 3.9e-91;
; Mismatches 0;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        HCV-1 NS3/4a mutant conformational antigen
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Location/Qualifiers 403
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              AAE18689 standard; Protein; 586 AA
                                                                                                                                                                                                                                                  .;
O
                                                                                                                                                                                                                        100.0%;
                                   Claim 3; Fig 9; 66pp; English
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               15-JUN-2000; 2000US-212082P.
02-APR-2001; 2001US-280811P.
02-APR-2001; 2001US-280867P.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     14-JUN-2001; 2001WO-US19369
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  (first entry)
                                                                                                                                                                                                                         Ouery Match 100.
Best Local Similarity 100.
Matches 182; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Hepatitis C virus type 1.
Synthetic.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Arcangel P,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              /note=
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                (CHIR ) CHIRON CORP
                                                                                                                                                                                                  182 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Misc-difference
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Misc-difference
                                                                                                                                                                                                                                                                                                                                                                                                                          RS 182
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      WO200196875-A2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                17-MAY-2002
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              20-DEC-2001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Chien DY,
                                                                                                                                                                                                  Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                          181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        AAE18689;
                                                                                                                                                                                                                                                                                                                          61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  윤
                                                                                                                                                                                                                                                                                                                        ç
                                                                                                                                                                                                                                                                                                                                               ag
                                                                                                                                                                                                                                                                                                                                                                         οχ
                                                                                                                                                                                                                                                                                                                                                                                                 a
                                                                                                                                                                                                                                                                                                                                                                                                                         ö
                                                                                                                                                                                                                                                                                                                                                                                                                                            g
```

```
61 GTRIIASPKGPVIQMYINVDKDLVGWPAPQGSRSLIPCTCGSSDLYLVTRHADVIPVRRR 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              61 GTRIIASPKGPVIQMYINVDQDLVGWPAPQGSRSLTPCTCGSSDLYLVTRHADVIPVRRR 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            The present invention relates to hepatitis C virus (HCV) core antigen and NS (nonstructural) 3,4a antibody combination assay that can detect both HCV antigens and antibodies present in a sample using a single solid matrix as well as immunoassay solid supports for use in the assay. The solid support is useful for detecting HCV infection in a biological sample. The present sequence is HCV-1 NS3/4a mutant conformational antigen. This sequence is used in the exemplification of the invention.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Hepatitis C virus; HCV; NS3/4a conformational epitope; seroconversion; immunoassay solid support; multiple epitope fusion antigen; MEFA; non-structural protein; mutant; mutein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1 MAPITAYAQQTRGLLGCIITSLTGRDKNQVEGEVQIVSTAAQTFLATCINGVCWTVYHGA
                                                                                                                                                                                                                                                                                                                                                                                                                                                Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Hepatitis C virus NS3/4a conformational epitope protein sequence
                                                            Immunoassay solid support useful for detecting hepatitis C virus infection in a biological sample, comprises at least one of HCV anti-core antibody and HCV NS3/4a epitope, bound to the support
                                                                                                                                                                                                                                                                                                                                                                                                                                             ö
                                                                                                                                                                                                                                                                                                                                                                                                     Length 686
                                                                                                                                                                                                                                                                                                                                                                                                 Score 946; DB 23; Length 6 Pred. No. 1.2e-89; 2; Mismatches 0; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                /note= "Wild-type Thr substituted by Pro"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             by Ile"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             substituted
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          *Wild-type Ser
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         AAU76377 standard; Protein; 686 AA.
                                                                                                                                              Example 2; Fig 4; 87pp; English
                                                                                                                                                                                                                                                                                                                                                                                                 Query Match
Best Local Similarity 98.9%;
Matches 180; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         14-JUN-2001; 2001WO-US19156
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  15-JUN-2000; 2000US-212082P.02-APR-2001; 2001US-280811P
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          /note-
2002-179522/23
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Misc-difference 403
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Misc-difference 404
                                                                                                                                                                                                                                                                                                                                                         686 AA;
```

ö

```
MEFA: anti-HCV antibody; NS3/4a conformational antigen;
HCV infection; mutant; mutein.
                                                                            Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Claim 2; Fig 3A-3D; 45pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  99.38;
                                                                                                                                                                                                                                                                         15-JUN-2000; 2000US-212082P. 02-APR-2001; 2001US-280811P. 02-APR-2001; 2001US-280867P.
                                                                                                                                                                                                                                                                                                                                                                   GEORGE-NASCIMENTO C. COIT D.
                                                                                                                                                                                                                                                14-JUN-2001; 2001US-0881654
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Matches 180; Conservative
                                    Hepatitis C virus type 1. Synthetic.
                                                                                                                                                                                                                                                                                                                                                                                                                       Arcangel P,
                                                                                                                                                                                                                                                                                                                                                                                              MEDI/) MEDINA-SELBY A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Query Match
Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                            WPI; 2003-147573/14.
N-PSDB; ABX14410.
                                                                                                                              Misc-difference 403
                                                                                                                                                       Misc-difference 404
                                                                                                                                                                                                                                                                                                                                         ARCANGEL P.
TANDESKE L.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            686 AA;
                                                                                                                                                                                                                                                                                                                              CHIEN D Y.
                                                                                                                                                                                                                                                                                                                                                                                                                                  Medina-Selby A;
                                                                                                                                                                                              US2002146685-A1
                                                                                                                                                                                                                        10-ocr-2002
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Seguence
                                                                                                                                                                                                                                                                                                                                                                                                                        Chien DY,
                                                                                                                                                                                                                                                                                                                                            ARCA/)
                                                                                                                                                                                                                                                                                                                              CHIE/
                                                                                                                                                                                                                                                                                                                                                                   GFOR/)
                                                                             Key
Reqion
 q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      og.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    The present invention relates to a new immunoassay solid support consisting essentially of at least one hepatitis C virus (HCV) NS3/4a conformational epitope and a multiple epitope fusion antigen (MEPA), bound to the support. The NS3/4a conformational epitope and/or MEPA reacts specifically with anti-HCV antibodies prosent in a biological sample from an HCV-infected individual. The immunoassay of the invention is useful for detecting hepatitis C virus infection in a biological sample. The method of the invention provides a sensitive, accurate diagnostic and prognostic tool to provide adequate patient care and to prevent transmission of HCV by blood and by blood products, or by
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          9
                                                                                                                                                                                                                                                                                                                                 personal contact. Use of NS3/4a conformational opinod products, or by personal contact. Use of NS3/4a conformational opinope in combination with MEFA, provides a sensitive and reliable method for detecting early HCV seroconversion. Use of MEFA has the added advantages of decreasing ansaking problems, improving sensitivity in detecting antibodies by allowing a greater number of epitopes on a unit surface area of substrate, and improving substrate, Detection accuracy is increased and the incidence of false results is reduced because of the identification and the use of highly immunogenic HCV antigons which are present during the early stages of HCV seroconversion. The present amino acid sequence represents the non-structural protein NS3/4a conformational epitope of the invention.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1 MAPITAYAQQTRGLLGC11TSLTGRUKNQVFGEVQIVSTAAQTFLATC1NGVCWTVYHGA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   GTRTIASPKGPVIQMYTNVDKDLVGWPAPQGSRSLTPCTCGSSDLYLVTRHADVIPVRRR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Gaps
                                                                                                                              Immunoassay solid support, useful for detecting hepatitis C virus infection in biological sample, comprises HCV NS3/4a conformational epitope and multiple epitope fusion antigen bound to the support \cdot
                                                    Coit D;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       0;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Immunoassay solid support; Hepatitis C Virus type-1; HCV-1;
NS3/4a conformational epitope; multiple epitope fusion antigen;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Score 946; DB 23; Length 686;
Pred. No. 1.2e-89;
2; Mismatches 0; Indels
                                                    George-nascicmento C,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 HCV-1 NS3/4a conformational antigen.
                                                      Tandeske L,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ABG72261 standard; Protein; 686 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ;
                                                                                                                                                                                   Claim 5; Fig 3; 92pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            99.38;
02-APR-2001; 2001US-280867P
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      06-MAR-2003 (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Query Match
Best Local Similarity 98.9
Matches 180; Conservative
                                                    Chien DY, Arcangel P, Medina-selby A:
                                                                                           WPI; 2002-090228/12.
                           (CHIR ) CHIRON CORP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     586 AA;
                                                                                                        N-PSDB; ABK15344
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          RS 182
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   182
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ABG72261;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           RESULT 4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           οg
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ద
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          õ
```

```
The present invention relates to immunoassays comprising Hepatitis C Virus (HCV) NS3/da conformational epitope and multiple epitope fusion antigen (MEFA), bound to a solid support. The NS3/da epitope and/or the multiple epitope fusion antigen react with anti-HCV antibodies present in a biological sample from an HCV-infected Individual. The immunoassays and methods of the invention are useful for detecting that it is support provides a sensitive and reliable method for detecting early HCV seroconversion. The assays can detect HCV infection acused by any C six known genotypes of HCV. The use of the multiple epitope fusion proteins decreases masking problems, improves sensitivity in detecting antibodies by allowing a greater number of epitopes on a unit area of substrate, and improves selectivity. The present sequence represents HCV-1) NS3/da conformational antigen, a mutant of the HCV-1 NS3/da polypeptide.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               61 GTRTIASPKGPVIQWYINVDQDLVGWPAPQGSRSLTPCTCGSSDLYLVTRHADVIPVRRR 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       61 GTRTIASPKGPVIQMYTNVDKDLVGWPAPQGSRSLTPCTCGSSDLYLVTRHADVIPVRRR 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Immunoassay solid support for detecting Hepatitis C Virus infection in biological samples, comprises Hepatitis C Virus conformational epitope and multiple epitope fusion antigen -
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1 MAPITAYAQQTRGLLGCIITSLTGRDKNQVEGEVQIVSTAAQTFLATCINGVCWTVYHGA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Gaps
/note "Corresponds to amino acid residues 1027-1711 of HCV-1 NS3/4m polypeptide"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Coit D;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                .;
0
                                                                                                                         Pro.
                                                                                                                                                                                             /note- "Substitution of wild-type Ser to Ile"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Length 686;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                0; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            George-Nascimento C,
                                                                                                                 /note= "Substitution of wild-type Thr to
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Score 946; DB 24;
Pred. No. 1.2e-89;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            2; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Tandeske L,
```

ö

QQ

g

ò

```
61 GTRIIASPKGPVIQMYTNVDKDLVGWPAPQGSRSLIPCTCGSSDLYLVTRHADVIPVRRR 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             GDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVCTRGVAKAVDFIPVESLETTM 180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         121 GDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVCTRGVAKAVDFIPVESLETTM 180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          20 LAPITAYAQQTRGLLGCIITSLTGRDKNQVEGEVQIVSTAAQTFLATCINGVCWTVYHGA 79
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1 MAPITAYAQQTRGLLGCIITSLTGRDKNQVEGEVQIVSTAAQTFLATCINGVCWTVYHGA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Peptide fragments of the non-structural protein (NS3) are reactive with and can detect antibodies against the NS3 domain of HCV. The peptides can be used for diagnosis of HCV infection. Nonspecific reaction can be inhibited and misdiagnosis of HCV infection can be
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ó
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  New peptide(s) reactive with anti-hepatitis C virus antibody for specific, early diagnosis of HCV infection
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Length 609;
                                                                                                                                                                                                                                                                      Peptide; antibody; hepatitis C virus; HCV; identification; diagnosis; non-A non-B hepatitis; NANB; detection.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   99.0%; Score 943; DB 15; 198.4%; Pred. No. 2.1e-89; iive 3; Mismatches 0;
                                                                                                                                                                                                                                             Hepatitis C virus non-structural protein 3.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Disclosure; Page 9-10; 15pp; Japanese.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          AAP92041
ID AAP92041 standard; protein; 1766 AA.
                                                                                                                                                        AAR51170 standard; peptide; 609 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                             (OLYU ) OLYMPUS OPTICAL CO LTD
                                                                                                                                                                                                                                                                                                                                                                                                                                 92JP-0209201.
                                                                                                                                                                                                                                                                                                                                                                                                    92JP-0209201.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              decreased. See AAR51162-70
                                                                                                                                                                                                                (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Query Match
Best Local Similarity 98.4
Matches 179; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       WPI; 1994-106803/13
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           609 AA;
                                                                                                                                                                                                                                                                                                                  Hepatitis C virus
                                                                          ||
| RS | 1207
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    RS 182
                                                       RS 182
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               201
                                                                                                                                                                                                                                                                                                                                                                                                     05-AUG-1992;
                                                                                                                                                                                                                                                                                                                                                                                                                                 05-AUG-1992;
                                                                                                                                                                                                                                                                                                                                                                          01-MAR-1994.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        - S2
                                                                                                                                                                                                                20-0CT-1994
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Sequence
121
                                                                                                                                                                                     AAR51170;
                                                       181
                                                                                                                              RESULT 6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             RESULT 7
                                                                                                                                         AAR51170
                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                         a
                                                       \tilde{c}
                                                                                 d
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ó
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            GTRT1ASPKGPV1QMYTNVIJKDLVGWPAPQGSRSIJTPCTCGSSDLYIJVTRHADVIPVRRR 120
                           121 GOSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVCTRGVAKAVDFJPVENLETTM 180
            GDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVCIRGVAKAVDFIPVESLEITM 180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1 MAPITAYAQQTRGLLGCIITSLTGRDKNQVEGEVQIVSTAAQTFLATCINGVCWTVYHGA 60
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  RNA was isolated from the plasma of a HCV seropositive human (designated "LG") and CDNA was prepared from it. The CDNA was PCR amplified using specific primers with sequences based on the prototype HCV-1 CDNA sequence (GENBANK M65321). Further amplification using nested primers resulted in 7 adjacent HCV DNA fragments which could be assembled into a foil-length sequence. The DNA sequence was determined and translated into the genomic amino acid sequence. Comparison of the LG genomic amino acid sequence with that from HCV-1 showed 134 amino acid differences.
                                                                                                                                                                                                                                                                        HCV genomic amino acid sequence isolated from infected human 13.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               New plasmid pHCV-162 is a mammalian expression systems for HCVT proteins – useful for diagnosing HCV infection and as vaccines for preventing HCV infection
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ó
                                                                                                                                                                                                                                                                                         Hepatitis C Virus; Non-A, non-B hepatitis Virus; HCV; NANBHV; Muman growth hormone; HGH; secretion signal; fusion protein; vaccine.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Length 3011;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 E
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Frail
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         .,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Score 944; DB 14;
Pred. No. 1.4c-88;
3; Mismatches 0;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (Updated on 25-MAR-2003 to correct PN field.)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Devare SG,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Example 1; Page 39-49; 100pp; English.
                                                                                                                                                                       AAR40120 standard; Protein; 3011 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Desai SM,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         99.1%;
ilarity 98.4%;
Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                              93WO-US00907
                                                                                                                                                                                                                                                                                                                                                                                                                                                                         92US-0830024
                                                                                                                                                                                                                                              (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Zeck BJ;
                                                                                                                                                                                                                                 (updated)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         WPI; 1993-258673/32.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Casey JM,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Sequence 3011 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (ABBO ) ABBOTT LAB.
                                                                                                                                                                                                                                                                                                                                                          Hepatitis C Virus.
                                                                                       ||
| RS 182
                                                                       RS 182
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Yamaguchi J,
                                                                                                                                                                                                                                                                                                                                                                                      W09315193-A1
                                                                                                                                                                                                                                                                                                                                                                                                                                              29-JAN-1993;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                        31-JAN-1992;
                                                                                                                                                                                                                              25-MAR-2003
                                                                                                                                                                                                                                             27-JAN-1994
                                                                                                                                                                                                                                                                                                                                                                                                                   05-AUG-1993
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      179;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            61
                                                                     181
                                                                                                                                                                                                  AAR40120;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Bode SL,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Matches
```

ó

9

g à 엄

```
AAP90158 standard; protein; 1786 AA
                                                                                                                                                                                                                            (CHIR ) CHIRON CORPORATION
                                      (updated)
(first entry)
                                                                                   Hepatitis C virus; vaccine
                                                                                                                                                                                                                                              Houghton M, Choo QL,
                                                                                                                                                                                                                                                                WPI; 1989-215054/30
                                                                                                                                                                                                                                                                                                                                                                                                         1786 AA;
                                                                                                                                                                                                                                                                          N-PSDB; AAN90327
                                                                                                      Pan troglodytes
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              RS 182
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               RS 491
                                                                                                                                                           18-NOV-1988;
                                                                                                                                                                              18-NOV-1987;
30-DEC-1987;
                                                                                                                                                                                                26-FEB-1988;
26-OCT-1988;
                                                                                                                        GB2212511-A.
                                     25-MAR-2003
10-NOV-1989
                                                                                                                                         26-JUL-1989
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                490
                                                                                                                                                                                                                                                                                                                                                                                                          Sequence
                   AAP90158
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            RESULT 9
AAP90164
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
ô
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     It is the sequence encoded in the open reading frame of hepatitis C virus CDNA inserts in clones 14i,m 11b, 7f, 7e, 8h, 33c, 40b, 37b, 35, 36, 81, 32, 33b, 25c, 14c, 8f, a33f, 33g and 39c. It is antipenic and could be used in immunoassay reagents and vaccines and to generate antibodies useful in diagnosis and passive immunotherapy for HCV infection/non-A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  61 GTRTIASPKGPVIQMYTNVDKDLVGWPAPQCSRSLTPCTCGSSDLYI.VTRHADVIPVRRR 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           9
                                                  Sequence encoded in the hepatitis C virus (HCV) cDNA inserts in clones 141, 11b, 7f, 7e, 8h, 33c, 40b, 37b, 35, 36, 8i, 32, 33b, 25c, 14c, 8f. 33f, 33g and 39c.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        310 LAPITAYAQQTRGLLCCIITSLTGRDKNQVEGKVQIVSIAAQTFLATCINGVCWTVYHGA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1 MAPITAYAQQTRGLLGCIITSLTGRDKNQVEGEVQIVSTAAQTFLATCINGVCWTVYHGA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Gaps
                                                                                                                                                                                                                                                                                                                                                  Claim 13; Figure 26-1, 26-2, 26-3, 26-4, 26-5, 26-6; 139pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ;
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Score 943; DB 10; Length 1766;
Pred. No. 8.8c-89;
3; Mismatches 0; Indels 0
                                                                                           Hepatitis C virus (HCV); non-A, non-B hepatitis (HANBH)
                                                                                                                                                                                                                                                                                                                               and associated nucleic acids and polypeptide(s)
                                                                                                                                                                                                                                                                                                                                                                                                                         (Updated on 25-MAR-2003 to correct PR field.) (Updated on 25-MAR-2003 to correct PI field.)
                                                                                                                                                                                     870S-0139886.
880S-0161072.
880S-0161072.
880S-0263584.
88US-0263584.
                                                                                                                                                                                                                                                                         Kuo G;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                        99.08;
                                                                                                                                                                    88EP-0310922
                            (updated)
(first entry)
                                                                                                                                                                                                                                                                                                                      Purified hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Local Similarity 98.4 tes 179; Conservative
                                                                                                                                                                                                                                                                        Houghton M, Choc QL,
                                                                                                                                                                                                                                                                                          WPI; 1989-159274/22
                                                                                                                                                                                                                                                       (CHIR ) CHIRON CORP
                                                                                                                                                                                                                                                                                                                                                                                                                                                     1766 AA;
                                                                                                              Hepatitis C virus.
                                                                                                                                                                                                                                                                                                     N-PSDB: AAN92097
                                                                                                                                                                                                                                                                                                                                                                                                                 non-B hepatitis.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      |-|
RS 491
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            RS 182
                                                                                                                                                                                                         26-FEB-1988;
06-MAY-1988;
26-OCT-1988;
                                                                                                                                                                     18-NOV-1988;
                                                                                                                                                                                       18-NOV-1987;
30-DEC-1987;
                           25-MAR-2003
02-MAR-1990
                                                                                                                                                 31-MAY-1989
                                                                                                                                                                                                                                    14-NOV-1988
                                                                                                                                EP318216-A
                                                                                                                                                                                                                                                                                                                                                                                                                                                       Sedneuce
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         430
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             490
        AAP92041;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          RESULT 8
g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
```

```
121 GDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVCTRGVAKAVDFIPVESLETTM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    The sequence is encoded by the composite cDNA of AAN90327. These antigens react with antibodies in patients with non-A non-B hepatitis (NANBH). They can be used to diagnose HCV-induced NANBH, to raise antibodies for immunoasasy or treatment, or to produce vaccines. (Updated on 25-MAR-2003 to correct PR [ield.)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  61 GTRTIASPKGPVIQMYTNVDKDLVGWPAPQGSRSLTPCTCGSSDLYLVTRHADVIPVRRR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1 MAPITAYAQQTRGLLGCIITSLTGRDKNQVEGEVQIVSTAAQTFLATCINGVCWTVYHGA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             prodn. of polynucleotide probes, dlagnosis, prevention and treatment
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Score 943; DB 10; Length 1786;
Pred. No. 8.9e-89;
3; Mismatches 0; Indels 0
Protein sequence of hepatitis c virus composite cDNA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      AAP90164 standard; protein; 2261 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Repatitis C virus gene - used for polypeptide(s) and antibodies for of infection.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Disclosure; fig 26; 30pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                             Kuo G;
                                                                                                                                                                                                                                                                                     87US-0122714.
87US-0139886.
88US-0161072.
88US-0263584.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Query Match 99.0%;
Best Local Similarity 98.4%;
Matches 179; Conservative
                                                                                                                                                                                                                                          88GB-0027024
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    25-MAR-2003 (updated)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         DXXXI
```

ö

Gaps

9

180

```
Hepatitis C virus (HCV); non-A, non-B hepatitis (HANBH)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Hepatitis C virus (HCV); non-A, non-B hepatitis (HANBH)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    AAP92050 standard; protein; 2436
                                                                                                                                                                                                                                                                                                                                                      ΰ
                                                                                                                                                                                 870S-0122714.
870S-0139886.
880S-0161072.
880S-0191263.
880S-0263584.
880S-0271450.
                                                                                                                                                 88EP-0310922
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            (updated)
(first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Query Match
Best Local Similarity 98.48
Matches 179; Conservative
                                                                                                                                                                                                                                                                                                                                                      Choo QL,
                                                                                                                                                                                                                                                                                                                                                                                       WPI; 1989-159274/22.
N-PSDB; AAN92103.
                                                                                                                                                                                                                                                                                                                   (CHIR ) CHIRON CORP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       2301 AA;
                                      Hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        K9-1 through 15e
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    RS 182
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     RS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          25-MAR-2003
02-MAR-1990
                                                                                                                                                                                                                                                          25-OCT-1988;
14-NCV-1988;
                                                                                                                                                                                                      30-DEC-1987;
26-FEB-1988;
                                                                                                                                               18-NOV-1988;
                                                                                                                                                                                    18-NOV-1987;
                                                                                                                                                                                                                                           06-MAY-1988;
                                                                                                                                                                                                                                                                                                                                                      Houghton M,
                                                                                                            31-MAY-1989
                                                                       EP318216-A
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     581
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       AAP92050
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 AAP92050
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     d
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ga
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       GTRTIASPKGPVIOMYTNVDKDLVGWPAPQGSRSLTPCTCGSSDLYLVTRHADVIPVRRR 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 GDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVCTRGVAKAVDFIPVESLETTM 180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Sequence encoded in the hepatitis C virus (HCV) cDNA inserts in clones 12f\ through\ 15c.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1 MAPITAYAQQTRGLLGCIITSLTGRDKNQVEGEVQIVSTAAQIFLATCINGVCWTVYHGA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         The sequence is the peptide encoded by the composite hepatitis C virus (HCV) CDNA of AAN90331. The polypeptides are used to diagnose HCV-induced NANBH, to raise antibodies for immunoassay or treatment, our to produce vaccines. (Updated on 25-MAR-2003 to correct PR field.)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Hepatitis C virus gene - used for prodn. of polynucleotide probes, polypeptide(s) and antibodies for diagnosis, prevention and treatment of infection.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       99.0%; Score 943; DB 10; Length 2261; ilarity 98.4%; Pred. No. 1.2e-88; Conservative 3; Mismatches 0; Indels 0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Indels
                                                                       Hepatitis C virus; clone 12f; clone 15e; probe; vaccine
                                      Peptide encoded by composite hepatitis C virus cDNA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Ź
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Disclosure; fig 32; 235pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  AAP92047 standard; protein; 2301
                                                                                                                                                                                                                                                        87US-0122714.
87US-0139886.
88US-0161072.
88US-0263584.
                                                                                                                                                                                                                       88GB-0027024
                                                                                                                                                                                                                                                                                                                                                      (CHIR ) CHIRON CORPORATION.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            (updated)
(first entry)
 (first entry)
                                                                                                                                                                                                                                                                                                                                                                                       Choo QL,
                                                                                                                                                                                                                                                                                                                                                                                                                            WPI; 1989-215054/30.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Local Similarity
les 179; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        2261 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                               N-PSDB; AAN90331
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     RS 182
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        582
                                                                                                              Pan troglodytes
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           - S
                                                                                                                                                                                                                       18-NOV-1988;
                                                                                                                                                                                                                                                            18-NOV-1987;
                                                                                                                                                                                                                                                                              30-DEC-1987;
26-FEB-1988;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            25-MAR-2003
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             02-MAR-1990
                                                                                                                                                                                                                                                                                                                 26-OCT-1988;
                                                                                                                                                                                                                                                                                                                                                                                         Houghton M,
01-NOV-1989
                                                                                                                                                 GB2212511-A
                                                                                                                                                                                  26-JUL-1989
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       AAP92047;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    401
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Seguence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Best Loca
Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 RESULT 10
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                AAP92047

ID AAP8

XX

XX

XC AAP9

DT 25-P

DT 02-P

XX

DE Sequence of the control of the cont
g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             QQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
```

```
It is the sequence encoded in the open reading frame of hepatitis C virus (HCV) cDNA inserts in clones 12f through 15e. It is antigenic and could be used in immunoassay reagents and vaccines and to generate antibodies useful in diagnosis and passive immunotherapy for HCV infection/non-A, non-B hepatitis.

(Updated on 25-MAR-2003 to correct PR field.)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Sequence encoded in the hepatitis C virus (HCV) cDNA inserts in clones
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      GDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVCTRGVAKAVDFIPVESLETTM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1 MAPITAYAQQTRGLLGCIITSLTGRDKNQVEGEVQIVSTAAQTFLATCINGVCWTVYHGA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ö
                                                                                                                                                                                                                                                                                                                                                                                                                          Length 2301;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Indels
Purified hepatitis C virus - and associated nucleic acids and polypeptide(s)
                                                                                                                                                                                                                                                                                                                                                                                                                    99.0%; Score 943; DB 10; 98.4%; Pred. No. 1.2e-88;
                                                                                Claim 13; Figure 32-1 - 32-7; 139 pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                             3; Mismatches
```

ö

```
61 GTRTIASPKGPVIQMYTNVDKDLVGWPAPQGSRSLTPCTCGSSDLYLVTRHADVIPVRRR 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 GDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVCTRGVAKAVDFIPVESLETTM 180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1 MAPITAYAQQTRGLLGCIITSLTGRDKNQVEGEVQIVSTAAQTFLATCINGVCWTVYHGA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          The sequence is the peptide encoded by the composite hepatitis C virus (HCV) cDNA of AAN90336. The polypeptides are used to diagnose HCV-induced NANBH, to raise antibodies for immunoassay or treatment, or to produce vaccines.

(N.B. This record was resubmitted to correct errors in the sequence.) (Updated on 25-MAR-2003 to correct PR field.)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Hepatitis C virus gene - used for prodn. of polynucleotide probes, polypeptide(s) and antibodies for diagnosis, prevention and treatment of infection.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Protein encoded by a cDNA compiled Hepatitis C virus cDNA clones.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Hepatitis C virus; HCV; antisense polynucleotide; polyprotein; viral infectivity; viral replication.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Query Match 99.0%; Score 943; DB 10; Length 2436; Best Local Similarity 98.4%; Pred. No. 1.3e-88; Matches 179; Conservative 3; Mismatches 0; Indels 0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Disclosure; fig 47-1 to 47-8; 235pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        AAB18540 standard; Protein; 2772 AA
                                                                                                                                                                                                                                     87US-0139886.
88US-0161072.
88US-0263584.
                                                                                                                                                                                                                                                                                                                                                                                                        Kuo G;
                                                                                                                                                                                                               87US-0122714
                                                                                                                                                                                                                                                                                                                                                (CHIR ) CHIRON CORPORATION.
                                                                                                                                                            88GB-0027024
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                15-JAN-2001 (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                  Houghton M, Choo QL,
                                                                                                                                                                                                                                                                                                                                                                                                                                                     WPI; 1989-215054/30.
N-PSDB; AAN90336.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   2436 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Hepatitis C virus
Pan troglodytes
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  181 RS 182
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       757
                                                                                                                                                                                                                                     30-DEC-1987;
26-FEB-1988;
26-OCT-1988;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              EP1034785-A2
                                                                                                                                                            18-NOV-1988;
                                                                                                                                                                                                               8-NOV-1987;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   - 8x
                                               GB2212511-A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    756
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          AAB18540;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     RESULT 13
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            AAB18540
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 $\circ{\circ}{\circ} \circ{\circ}{\circ} \circ{\circ}{\circ} \circ{\circ}{\circ} \circ{\circ}{\circ} \circ{\circ}{\circ} \circ{\circ}{\circ} \circ \ci
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          It is the sequence encoded in the open reading frame of hepatitis C virus (HCV) cDNA inserts in clones K9-1 through 15e. It is antigenic and could be used in immunoassay reagents and vaccines and to generate antibodies useful in diagnosis and passive immunotherapy for HCV infection/non-A, non-B hepatitis.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              635
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          GTRTIASPKGPVIQMYTNVDKDLVGWPAPQGSRSLTPCTCGSSDLYLVTHHADVIPVRRR 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      GDSRGSLLSPRPISYLKGSSGGPLLCPACHAVGIFRAAVCTRGVAKAVDFIPVESLETTM 180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1 MAPITAYAQQTRGLLGCIITSLTGRDKNQVEGEVQIVSTAAQTFLATCINGVCWTVYHGA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Caps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    0;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         99.0%; Score 943; DB 10; Length 2436;
98.4%; Pred. No. 1.3e-88;
Live 3; Mismatches 0; indels 0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Hepatitis C virus; clone 15e; clone k9-1; probe; vaccine.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Purified hepatitis C virus - and associated nucleic acids and polypeptide(s)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Peptide encoded by composite hepatitis C cDNA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Claim 13; Figure 47-1 - 47-8; 139 pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (Updated on 25-MAR-2003 to correct PR field.) (Updated on 25-MAR-2003 to correct PI field.)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 AAP90288 standard; protein; 2436 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Kuo G;
                                                                                                                                                                                                                          87US-0122714.
87US-0139886.
88US-0161072.
88US-0191263.
88US-0263584.
88US-0271450.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        (updated)
(updated)
(first entry)
                                                                                                                                                                                   88EP-0310922
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Query Match

Best Local Similarity 98.45
Matches 179; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Houghton M. Choo QL,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                WPI; 1989-159274/22.
                                                                                                                                                                                                                                                                                                                                                                                                                         (CHIR ) CHIRON CORP.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         2436 AA;
                   Hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          N-PSDB; AAN92106
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           RS 182
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          RS 757
                                                                                                                                                                                                                          18-NOV-1987;
30-DEC-1987;
26-FEB-1988;
06-MAY-1988;
26-OCT-1988;
14-NOV-1988;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      25-MAR-2003
19-JUL-2001
01-NOV-1989
                                                                                                                                                                                18-NOV-1988;
                                                                                                                              31-MAY-1989
                                                                        EP318216-A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      AAP90288;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                989
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            RESULT 12
```

Dp

qq ò

ð

g à

õ

ö

635

œ

```
gp53(BVDV)/gp55; hog cholera virus; pestivirus; NS1; flavivirus
                                                                                                                                                                                                                                                                                                                                                                                                       99.0%; Score 943; DB 14;
98.4%; Pred. No. 1.6e-88;
live 3; Mismatches 0;
                                                                                                                                                                                                                                                                                                                                                          (Updated on 25-MAR-2003 to correct PN field.)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Composite hepatitis C virus (HC-J1/CDC/CHI).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   AAR70230 standard; protein; 2894 AA.
                                                                                                                                                                                                                                                              Disclosure; Fig 9; 106pp; English.
                                                                                            92WO-US07683
                                                                                                                   9105-0759575
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (updated)
(first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                     Best_Local Similarity 98.48
Matches 179; Conservative
                                                                                                                                                                  Weiner AJ;
                                                                                                                                                                                        WPI; 1993-117468/14.
                                                                                                                                         (CHIR ) CHIRON CORP
                                                                                                                                                                                                                                                                                                                                                                                  2816 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Hepatitis C virus
                     Hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    RS 1207
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  RS 182
                                             W09306126-A1
                                                                                            11-SEP-1992;
                                                                                                                   13-SEP-1991;
                                                                                                                                                                                                                                      HCV isolates
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 25-MAR-2003
07-NOV-1995
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         EP644202-A1.
                                                                                                                                                                  Houghton M,
                                                                     01-APR-1993
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  181
                                                                                                                                                                                                                                                                                                                                                                                    Sequence
                                                                                                                                                                                                                                                                                                                                                                                                        Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          AAR70230;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        AAR70230
qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             971
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  121 GDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVCTRGVAKAVDFIPVESLETTM 180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      9
                                                                                                                                                                                                                                                          The specification describes a pharmaceutical composition which comprises a hepatitis C virus (HCV) antisense polynuclectide. The HCV is characterized by a positive stranded RNA genome which has 40% homology at the polypectide level to a HCV polyprotein. The antisense polynuclectide binds to cellular polynuclectides which enhance and/or are required for viral infectivity, replicative ability or chronicity. The antisense polynuclectides may also be designed to bind with high specificity, to be of increased stability, to be stable and to have low toxicity. The composition also comprises an agent which causes viral RNA to be inactive. The composition is used for preventing HCV replication in a system. The present sequence is encoded by a novel HCV cDNA sequence, which is used in the course of the invention.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    GTRTIASPKGPVIQMYTNVDKDLVGWPAPQGSRSLTPCTCGSSDLYLVTRHADVIPVRRR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                            0; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Polymerase chain reaction; PCR; amplify; primer; hepatitis C virus; HCV; asymptomatic; chronically infected; cpitope; viral isolate; domain; immunological; cross-reactive; envelope protein; vaccine;
                                                                                                                                                                                       Novel composition comprising a hepatitis C virus antisense polynucleotide which is complementary to or corresponds to a sense strand of the virus genome, and selectively hybridises to it.
                                                                                                                                                                                                                                                                                                                                                                                                                                                     Match 99.0%; Score 943; DB 21; Length 2772; Local Similarity 98.4%; Pred. No. 1.6e-88; es 179; Conservative 3; Mismatches 0; Indels 0.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  AAR34009 standard; Protein; 2816 AA.
                                                                                                                                                                                                                                      Example; Fig 16; 75pp; English.
                                            89US-0325338.
89US-0341334.
89US-0355002.
90EP-0302866.
                                                                                                                               Kuo G;
                      2000EP-0109602
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (updated)
(first entry)
                                                                                                                               Choo 0,
                                                                                                                                                      WPI; 2000-566891/53.
                                                                                                       (CHIR ) CHIRON CORP
                                                                                                                                                                                                                                                                                                                                                                                                                                2772 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    HCV-1 polyprotein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ||
|RS || 1093
                                                                                                                                                                N-PSDB; AAA75296
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 RS 182
                      16-MAR-1990;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 25-MAR-2003
26-JUL-1993
                                                        20-APR-1989;
18-MAY-1989;
                                                                                 16-MAR-1990;
                                                                                                                               Houghton M,
                                             17-MAR-1989
13-SEP-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         AAR34009;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 181
                                                                                                                                                                                                                                                                                                                                                                                                                                  Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                        Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          RESULT 14
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      AAR34009
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     QQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
```

```
1086 GTRTIASPKGPVIQMYTNVDQDLVGWPAPQGSRSLTPCTCGSSDLYLVTRHADVIPVRRR 1145
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            121 GDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVCTRGVAKAVDFIPVESLETTM 180
                                                                                                                                                                                                                    This sequence represents the entire hepatitis C virus polyprotein. HCV is a member of the flavivirus family and appears to encode a basic polypeptide domain ("C") at the N-terminal of the viral polyprotein, followed by two glycoprotein domains ("E1", "E2/NS1"), upstream of the nonstructural genes NS2 through NS5. See also AAQ39134-48, AAR33982-4008 and AAR38088-89
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               GTRITASPKGPVIQMYINVDKDLVGWPAPQGSRSLTPCTCGSSDLYLVTRHADVIPVRRR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1 MAPITAYAQQTRGLLGCIITSLTGRDKNQVEGEVQIVSTAAQTFLATCINGVCWTVYHGA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Gaps
Immuno-reactive hepatitis C virus polypeptide compsns. - contg. at least 2 sequences from the first variable domain of distinct
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ó;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Composite hepatitis C virus; HC.31/CDC/CHI; HCV; non-A non-B; synthetic antigens; blood screening.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Length 2816;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Indels
```

ö

22-MAR-1995

```
ô
                                                                                                                                                                                                                                                                                                                                                                                                                  1 MAPITAYAQQTRGLLGCIITSLIGRDKNOVEGEVOIVSTAAQTFLATCINGVCWTVYHGA 50
                                                                                                                                                                                                                                                                                                                                                                                             Gaps
                                                                                                                                                                                                                              AAR70230 is the composite hepatitis C virus (HC-J1/CDC/CHI) protein from which the synthetic HCV antigens described in AAR70210-R70229 were derived. These synthetic antigens can be used to screen blood, or blood products for the presence HCV, they can also be used in various specific assays for the detection of HCV antibodies, and antigens, or as immunogens.

(Updated on 25-MAR-2003 to correct PV field.)
                                                                                                                                                       Synthetic antigens for the detection of hepatitis C virus antibodies \cdot comprise portions of the HCV peptide sequence, for use in screening blood and blood products
                                                                                                                                                                                                                                                                                                                                                                                            .;
0
                                                                                                                                                                                                                                                                                                                                                                  99.0%; Score 943; DB 16; Length 2894; 98.4%; Pred. No. 1.7e-88;
                                                                                                                                                                                                                                                                                                                                                                                            Indels
                                                                                                            Pollet D. Van Heuverswyn H;
                                                                                                                                                                                                                                                                                                                                                                                            Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Search completed: August 30, 2003, 19:12:19 Job time : 42.2251 secs
                                                                                                                                                                                                          Disclosure; Fig 1; 51pp; English.
                                                                                                                                                                                                                                                                                                                                                                                           3;
                                                90EP-0124241.
                         94EP-0108611.
                                                                                                                                                                                                                                                                                                                                                                  Query Match
Best Local Similarity 98.4*
Matches 179; Conservative
                                                                                    (INNO-) INNOGENETICS NV
                                                                                                            Maertens G,
                                                                                                                                    WPI; 1995-116946/16.
                                                                                                                                                                                                                                                                                                                                             2894 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        RS 1207
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  RS 182
                                                14-DEC-1990;
14-DEC-1990;
                         14-DEC-1990;
                                                                                                            Deleys RJ,
                                                                                                                                                                                                                                                                                                                                             Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  181
ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                       đ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ολ
```

```
Sense strand of the Hepstitis C virus Sense strand of HC Compiled HCV cDNA. Hepstitis C virus HCV polyprotein co HCV polyprotein co Hepstitis C virus Hepstitis C virus Hepstitis C virus Hepstitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                            Nucleotide sequency Hepatitis C virus Hepatitis C virus Hepatitis C virus Hepatitis C virus And encoding hepatitis C virus Hepatitis C virus Hepatitis C virus Nucleotide sequency Plasm.d pTSH4./HIV HCV NS3 protease C pNB182delt44 HT e HCV Solubilised NS PROME HCV NS3 DNA. Hepa
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Hepalitis C virus
Hepalitis C virus
Hepalitis C virus
Hepalitis C virus
CDNA encoding hepa
                                                                                                                             HCV-1 NS3/4a mutan
                                                                                                                                                                          Combined open read
Hepatitis C virus
Combined open read
                                                                                                                                                                                                                                         Composite hepatiti
Hepatitis C virus
                                                                                                                                                                                                                                                                       CDNA sequence comp
Nucleotide sequenc
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Plasmid pET-BS(+)/
                                                                                                                                                            DNA encoding HCV-1
                                                                                                          Hepatitis C virus
                                                                                                                                             Hepatitis C virus
                                                                                                                                                                                                                         Combined open
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Hepatitis, NS3 protease; viral replication; chronic liver disease;
liver failure; liver cancer; ds.
and is derived by analysis of the total score distribution.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Hepatitis C virus NS3 protease coding sequence.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ALIGNMENTS
                               SUMMARIES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       AAT42394
AAT58401
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          4AT38903
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     AAX23259
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ABK87300
                                                                                                                                                                                                                                                                                                                                                                                                                                                             AAX26737
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ABK87286
                                                                                                                                                                                                                                                                                                                        AA01056
                                                                                                                                                                                                                                                                                                                                                                                                                                                AAD31767
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ABK87
                                                                                                                                                                                                                                                                                                                                        ZAZ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                BP
                                                                                                            AAA70344 standard; DNA; 549
                                                                              В
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              (first entry)
                                             %
Query
Match Length D
                                                                                                                                                                                                                                                                                                                                                                                                                                                             9185
9185
8316
8316
9502
9646
12980
12980
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            depatitis C virus.
                                                                                                                           19-DEC-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          \phi
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             AAA70344;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Rey
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                RESULT 1
                                                                                                                                                                                                                       Result
                                                                              ç
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                "MODEL-frame+_pol.model - DEV-xlp
-Q-/cgn2_1/USPTO_spool/US0996594/runat_29082003_i51918_28302/app_query.fasta_i.2872
-Q-/cgn2_1/USPTO_spool/US09965594/runat_29082003_i51918_28302/app_query.fasta_i.2872
-DB-N_Genseq_195un03 -QFMT-i.exp-suprix-rng -MINMATCH-0.1 - LOOPECL-0
-LUSPEXT-0 -(NNISS-bits - START-1 - END--1 -MATRIX-blos.m62 - TRANS-human40.cdi
-LIST-45 -DOCALIGN-200 - THR_SCORE-pct - THR_MAX*100 - THR_MIA-0 - ALIGN-15
-MODE-LCCAL -OUTFMT-pto -NORM-ext -HEAPSIZE-56C -MINLEN-0 -MAXLEN-200000000
-USER-US09965594_eCGN_1_1.112.erunat_29082003.151918_28302 - NCPU-6 - ICPU-3
-DOCAMAP - LARGEQUERY - NEG_SCORE-0 - WAI T - DSPRILCCK-100 - LONGLOG
-DEV_TIMEOUT-120 -WARN_IIMEOUT-80 - THREADS-1 - XGAPOP=10 - XGAPDFR=0, S - FGAPOP=6
-FGAPEXT-7 - YGAPOP=10 - YGAPEXT-0.5 - DELEXI-7
                                                                                                    August 30, 2003, 19:13:57; Search time 169.009 Seconds (without alignments) 2906.924 Million cell updates/sec
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             N_cenescq_19Jun03:+

| S1DS1/gcqdata/genescqn-embl/NA198C_DAT:+
| S1DS1/gcqdata/genescqn-embl/NA1981_DAT:+
| S1DS1/gcqdata/genescqn-embl/NA1981_DAT:+
| S1DS1/gcqdata/genescqn-embl/NA1982_DAT:+
| S1DS1/gcqdata/genescqn-embl/NA1983_DAT:+
| S1DS1/gcqdata/genescqn-embl/NA1985_DAT:+
| S1DS1/gcqdata/genescqn-embl/NA1985_DAT:+
| S1DS1/gcqdata/genescqn-embl/NA1985_DAT:+
| S1DS1/gcqdata/genescqn-embl/NA1985_DAT:+
| S1DS1/gcqdata/genescqn-embl/NA1985_DAT:+
| S1DS1/gcqdata/genescqn-embl/NA1990_DAT:+
| S1DS1/gcqdata/genescqn-embl/NA1990_DAT:+
| S1DS1/gcqdata/genescqn-embl/NA1990_DAT:+
| S1DS1/gcqdata/genescqn-embl/NA1990_DAT:+
| S1DS1/gcqdata/genescqn-embl/NA1990_DAT:+
| S1DS1/gcgdata/genescqn-embl/NA1990_DAT:+
| S1DS1/gcgdata/genescqn-embl/NA2001_DAT:+
| S1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed,
                                                                                                                                                                                                     1 MAPITAYAQQTRGLLGCIIT.......GVAKAVDFIPVESLETIMRS 182
                                                                                                                                                                                                                                                                                                                                                                5105512
           GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.
                                                                         nucleic search, using frame_plus_p2n model
                                                                                                                                                                                                                                                                                                                                                            Total number of hits satisfying chosen parameters:
                                                                                                                                                                                                                                                                                                                                2552756 segs, 1349719017 residues
                                                                                                                                                                                                                                                                                                                                                                                                                                         Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries
                                                                                                                                                                                                                                                   0.5
                                                                                                                                                                                                                          BLOSUM62
Xgapop 10.0 , Xgapext C
Ygapop 10.0 , Ygapext (
Fgapop 6.0 , Fgapext )
Delop 6.0 , Delext
                                                                                                                                                                                                                                                                                                                                                                                           Minimum DB seq length: 0
Maximum DB seq length: 2000000000
                                                                                                                                                                 US-09-965-594-1
953
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Command line parameters:
                                                                                                                                                                     Title:
Perfect score:
                                                                                                                                                                                                                                    Scoring table:
                                                                         protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Database :
                                                                                                                                                                                                     sedneuce:
                                                                                                                                                                                                                                                                                                                                  Searched:
                                                                                                       Run on:
```

```
121
                                                                                                                                                                         RESULT 2
                                                                                                                                                                                               q
                                        ò
                                                              g
                                                                                     ð
                                                                                                       g
                                                                                                                             ò
                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          240
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    80
                                                                                                                                                                                                                                                                     The present sequence is the coding sequence for the Hepatitis C virus (HCV) NS3 protease enzyme. This protein is essential for the replication of the virus, acting to cleave its replicative proteins from the polyprotein produced from the HCV genome. NS4A is also needed for this process and inhibitors of the two proteins should act as antiviral treatments of HCV infection. This is useful'as HCV can lead to chronic liver disease such as cirrhosis, liver failure and liver cancer. The present invention concerns a number of NS3 mutants and NS3-NS4A fusion proteins which can be used to identify inhibitors of this type, as well.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      20
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          9
                                                                                                                                                                                           Modified hepatitis C virus (HCV) NS3 protease comprising at least 1 substitution of a hydrophobic alpha-helix 0 amino acid to a hydrophilic amino acid, useful for screening inhibitors that may treat hepatitis {\sf C}
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SerLeuThrGlyArgAspLysAsnGlnValGluGlyGluValGlnIleValSerThrAla
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         41 AlaGlnThrPheLeuAlaThrCysIleAsnGlyValCysTrpThrValTyrHisGlyAla
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    GlyThrargThrIleAlaSerProLySGlyProValIleGlnMetTyrThrAsnValAsp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           GlySerScrAspLeuTyrLeuValThrArgHisAlaAspVallleProValArgArgArg
                                                                                                                                                                                                                                                                                                                                                                                                                                     549
182
0
0
0
                                                                                                                                                                                                                                                                                                                                                                                                       Sequence 549 BP; 83 A; 180 C; 142 G; 144 T; 0 other;
                                                                                                                                                                                                                                                                                                                                                                                                                                    Length:
Matches:
Conservative:
Mismatches:
Indels:
Gaps:
                                                                                                                                         Goldfarb
           protease"
                                                                                                                                        Zhang Y,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                US-09-965-594-1 (1-182) x AAA70344 (1-549)
                                                                                                                                                                                                                                                   Example 1: Fig 9; 65pp; English.
                                                                                                                    8
          "NS3
                                                                                                                                                                                                                                                                                                                                                                                                                                    6.35e-80
953.00
100.00%
100.00%
                                                                                                                    SQUIBB
                                                                                              99US-0115271.
                                                                                                                                        Weinheimer S,
                                                                         2000WO-US00345
          /product=
                                                                                                                    (BRIM ) BRISTOL-MYERS
                                                                                                                                                              WPI; 2000-465976/40.
                                                                                                                                                                                                                                                                                                                                                                                                                                                           Percent Similarity:
Best Local Similarity:
                                                                                                                                                                       P-PSDB; AAB15211
                                WO200040707-A1
                                                                                                                                         Wittekind M,
                                                                         06-JAN-2000;
                                                                                              08-JAN-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                           Alignment Scores:
Pred. No.:
                                                                                                                                                                                                                                                                                                                                                                                    complexes.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                21
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          101
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Match:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Query
DB:
                                                                                                                                                                                                                                                                                                                                                                                                                                                  Score:
```

qq

ò g ò g ò a ò g

ò

```
420
                                                                                              160
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               The present invention relates to hepatitis C virus (HCV) core antigen and NS (nonstructural) 3/4a antibody combination assay that can detect both HCV antigens and antibodies present in a sample using a single solid matrix as well as immunoassay solid supports for use in the assay. The solid support is useful for detecting HCV infection in a biological sample. The present sequence is a DNA encoding HCV-1 NSJ/4a mutant conformational antigen. This sequence is used in the exemplification
                                                                                                                                                                                        GlyAspSerArgGlySerLeuLeuSerProArgProlleSerTyrLeuLysGlySerSer
                         GGTGACTCCCGTGGTTCCCTGCTGTGTCCCCGCGTCCGATCTCCTACCTGAAAGGTTCCTCC
                                                                                              GlyGlyProLeuLeuCysProAlaGlyHisAlaValGlyIlePheArgAlaAlaValCys
                                                                                                                      ä
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Coit
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Immunoassay solid support useful for detecting hepatitis C virus infection in a biological sample, comprises at least one of HCV anti-core antibody and HCV NS3/4a epitope, bound to the support
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           "HCV-1 NS3/4a conformational antigen"
"CDS does not include stop codon"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ds.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ပဲ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Hepatitis C virus; NS3/4a antigen; HCV infection; mutant;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          George-Nasciemento
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Seguence 2058 BP; 419 A; 634 C; 580 G; 425 T; 0 other;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  HCV-1 NS3/4a mutant conformational antigen encoding
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Tandeske L,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Location/Qualifiers
1..686
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Example 2; Fig 4; 87pp; English
                                                                                                                                                                                                                                                                                                                                                                                                                                       ВР
                                                                                                                                                                                                                                                                                                                                                                                                                                       AAD29795 standard; DNA; 2058
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        15-JJN-2000; 2000US-212082P. 02-APR-2001; 2001US-280811P. 02-APR-2001; 2001US-280867P.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         14-JUN-2001; 2001WO-US19369.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         /*tag= a
/product=
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        /note= "C
/partial
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Hepatitis C virus type 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Arcangel P,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             WPI; 2002-179522/23.
P-PSDB; AAE18689.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (CHIR ) CHIRON CORP
                                                                                                                                                                                                                                                                                          181 ArgSer 182
                                                                                                                                                                                                                                                                                                                           CGTICC 546
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              the invention.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Medina-Selby A;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                WO200196875-A2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      17-MAY-2002
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           20-DEC-2001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Synthetic
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Chien DY,
                                                                                              141
                                                                                                                                             421
                                                                                                                                                                                                                                                                                                                                        541
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        AAD29795;
                                              361
```

ä

3

ò g ò g à g ò ö a ò qq ô g ò

```
The present invention relates to a new immunoassay solid support consisting essentially of at least one hepatitis C virus (HCV) NS34a conformational epitope and a multiple epitope fusion antigen (MEPA).

Conformational epitope and a multiple epitope fusion antigen (MEPA).

Bound to the support. The NS3/4a conformational epitope and/or MEPA reacts specifically with anti-HCV antibodies present in a biological sample from an HCV-infected individual. The immunoassay of the invention is useful for detecting hepatitis C virus infection in a biological sample. The method of the invention provides a sensitive, accurate prevent transmission of HCV by blood and by blood products, or by prevent transmission of HCV by blood and by blood products, or by personal contact. Use of NS3/4a conformational epitope in combination with MEPA, provides a sensitive and reliable method for detecting early HCV seroconversion. Use of MEPA has the added advantages of decreasing masking problems, improving sensitivity in detecting antibodies by allowing a greater number of epitopes on a unit surface area of substrate, and improving substrate. Detection accuracy is increased and
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              the incidence of false results is reduced because of the identification and the use of highly immunogenic HCV antigens which are present during the early stages of HCV seroconversity the present nucleic acid sequence encodes the non-structural protein NS3/4a conformational epitope of the
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     61 AGCCTAACTGGCCGGGACAAAACCAAGTGGAGGGTGAGGTCCAGATTGTGTCAACTGCT 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SerLeuThrGlyArgAspLysAsnGlnValGluGlyGluValGlnIleValSerThrAla
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   41 AlaGInThrPheLeuAlaThrCysIleAsnGlyValCysTrpThrValTyrHisGlyAla
                                                                                                                                                                                                                                                                  Immunoassay solid support, useful for detecting hepatitis C virus infection in biological sample, comprises HCV NS3/4a conformational epitope and multiple epitope fusion antigen bound to the support \cdot
                                                                                                                                                                              Coit
                                                                                                                                                                              George-nasciemento C,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             2058
1180
2
0
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Sequence 2058 BP; 419 A; 633 C; 581 G; 425 T; 0 other;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Matches:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Length:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   US-09-965-594-1 (1-182) x ABK15344 (1-2058)
                                                                                                                                                                              Tandeske L,
                                                                                                                                                                                                                                                                                                                            Disclosure; Fig 3; 92pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1.46e-78
                                                                                     15-JUN-2000; 2000US-212082P.
02-APR-2001; 2001US-280811P.
02-APR-2001; 2001US-280867P.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           946.00
100.00%
98.90%
99.27%
                                                           14-JUN-2001; 2001WO-US19156
                                                                                                                                                                              Arcangel P,
                                                                                                                                                                                                                        WPI; 2002-090228/12.
P-PSDB; AAU76377.
                                                                                                                                                (CHIR ) CHIRON CORP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Percent Similarity:
Best Local Similarity:
                                                                                                                                                                                            Medina-selby A;
WO200196870-A2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Alignment Scores:
                               20-DEC-2001
                                                                                                                                                                              Chien DY,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          invention
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         21
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    a
   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   á
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       GlyAspSerArgGlySerLeuLeuSerProArgProlieSerTyrLeuLysGlySerSer 140
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ThrargGlyValAlaLysAlaValAspPheIleProValGluSerLeuGluThrThrMet 180
                                                                                                                                                                                                                                                                 GlySerSerAspLeuTyrLeuValThrArgHisAlaAspValIleProValArgArg 120
                                                                                                                                                                              9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Hepatitis C virus; HCV; NS3/4a conformational epitope; seroconversion; immunoassay solid support; multiple epitope fusion antigen; MEFA;
                                                                                                                                                MetalaprollethralaTyralaGlnGlnThrargGlyLeuLeuGlyCysIleIleThr
                                                                                                                                                                              ATGGCGCCCATCACGCCTACGCCCAGCAGGGGCCTCCTAGGGTGCATAATCACC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              /partial
/product- "HCV NS3/4a conformational epitope"
/note- "This sequence lacks a stop codon"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Hepatitis C virus NS3/4a conformational epitope gene seguence
              2058
180
2
0
0
              Length:
Matches:
Conservative:
Mismatches:
Indels:
                                                                                          Gaps:
                                                                                                                      US-09-965-594-1 (1-182) x AAD29795 (1-2058)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           non-structural protein; gene; ds
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           BP.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ABK15344 standard; DNA; 2058
                             946.00
100.00%
98.90%
99.27%
24
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   08-MAY-2002 (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Þ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  2058
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  /*tag=
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ArgSer 182
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     111111
541 AGGTCC 546
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Hepatitis C virus.
                                                            Best Local Similarity:
                                                Percent Similarity:
Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        481
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ABK15344;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                           301
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     361
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          161
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    181
                                                                                                                                                                                                                                                                                                                                                                                                                                              101
                                                                            Query Match:
DB:
                 ..
No.:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            RESULT 3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             a
                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
```

ò

9

```
140
                                             300
                                                                                                                                                                                                                          420
                                                                                                                                                                                                                                                                    091
                                                                                                                                                                                                                                                                                                               480
                                                                                                                                                                                                                                                                                                                                                            180
                                                                                       LysAspLeuValGlyTrpProAlaProGlnGlySerArgSerLeuThrProCysThrCys
                                                                                                                                                                               GlyAspScrArgClyScrLeuLeuScrProArgProllcScrTyrLcuLysGlySerScr
                                                                                                                                                                                                    Coit
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Immunoassay solid support; Hepatitis C Virus type-1; HCV-1; NS3/4a conformational epitope; multiple epitope fusion antigen; MEFA; anti-HCV antibody; NS3/4a conformational antigen; HCV infection; mutant; gene; ds.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ú
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   /product- "NS3/4a conformational antigen"
/note- "This sequence lacks a stop codon"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          George-Nascimento
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 encoding HCV-1 NS3/4a conformational antigen.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Location/Qualifiers
1..2058
/*tag= a
/partial
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ŗ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Tandeske
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ВР
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ABX14410 standard; DNA; 2058
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 15-JUN-2000; 2000US-212082P.
02-APR-2001; 2001US-280811P.
02-APR-2001; 2001US-280867P.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ú
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           2001US-0881654
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         06-MAR-2003 (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  GEORGE-NASCIMENTO
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Hepatitis C virus type 1. Synthetic.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Arcangel P,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (COIT/) COIT D. (MEDI/) MEDINA-SELBY A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CHIEN D Y.
ARCANGEL P.
TANDESKE L.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          WPI; 2003-147573/14.
P-PSDB; ABG72261.
                                                                                                                                                                                                                                                                                                                                                                                                                                               181 ArgSer 182
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          541 AGGTCC 546
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        US2002146685-A1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Medina-Selby A;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           14-JUN-2001;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                10-0CT-2002
                                             241
                                                                                                                                                                             121
                                                                                                                                                                                                                       361
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ABX14410;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           (CHIE/) (ARCA/) A (TAND/) 1 (GEOR/) G
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Chien DY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ABX1410
ABX1410
ABX1410
ABX14
AAC ABXX
AC ABXX
AX AX AX AX
AX AX AX AX
AX AX AX
AX AX AX
AX AX AX
AX AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
AX AX
A
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ONA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                a
                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                      Q
                                                                                                                                                                 οy
                                                                                                                                                                                                           QQ
                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                   Ob
                                                                                                                                                                                                                                                                                                                                           5
                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                 q
                                                                                ŏ
```

```
The present invention relates to immunoassays comprising Hepatitis C Virus (HCV) NS3/da conformational epitope and multiple epitope fusion antigon react with anti-HCV antibodies antigon (MEPA), bound to a solid support. The NS3/da epitope and/or the multiple epitope fusion antigen react with anti-HCV antibodies present in a biological sample from an HCV infected individual. The immunoassays and methods of the invention are useful for detecting HCV infection in a biological sample. The inventive immunoassay soils support provides a sensitive and reliable method for detecting early HCV seroconversion. The assays can detect HCV infection caused by any HCV seroconversion. The assays can detect HCV infection caused by any proteins decreases masking problems, improves sensitivity in detecting antibodies by allowing a greater number of epitopes on a unit area of substrate, and improves selectivity. The present sequence encodes HCV type I (HCV-I) NS3/da conformational antigen, a mutant of
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          240
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       100
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 300
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           360
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        140
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    420
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   160
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              80
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           40
   Hepatitis C Virus infection in C Virus conformational epitope
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            61 AGCCTAACTGGCCGGGGACAAAACCAAGTGGAGGGTGAGGTCCAGATTGTGTCAACTGCT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    41 AlaGlnThrPheLeuAlaThrCysIleAsnGlyValCysTrpThrValTyrHisGlyAla
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  GlyGlyProLeuLeuCysProAlaGlyHisAlaValGlyIlePheArgAlaAlaValCys
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   MetalaProlleThralaTyralaGlnGlnThrArgGlyLeuLeuGlyCysIleIleThr
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             21 SerbeuThrGlyArgAspLysAsnGlnValGluGlyGluValGlnIleValSerThrAla
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                121 GCCCAAACCTTCCTGGCAACGTGCATCAATGGGGTGTGCTGGACTGTCTACCACGGGGCC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ThrargGlyValalaLysAlaValAspPheIleProValGluSerLeuGluThrThrMet
                                                                                                                                                                                                                                                                                                                          Sequence 2058 BP; 419 A; 633 C; 581 G; 425 T; 0 other;
                                                                                                                                                                                                                                                                                                                                                                    2058
180
2
0
0
                                                                                                                                                                                                                                                                                                                                                                                    Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                      Length:
Immunoassay solid support for detecting
biological samples, comprises Hepatitis
                                                                                                                                                                                                                                                                                                                                                                                                                                                                         US-09-965-594-1 (1-182) x ABX14410 (1-2058)
                 biological samples, comprises Hepatin.
and multiple epitope fusion antigen
                                                              Disclosure; Fig 3A-3D; 45pp; English
                                                                                                                                                                                                                                                                                             the HCV-1 NS3/4a polypeptide
                                                                                                                                                                                                                                                                                                                                                                   1.46e-78
946.00
100.00%
98.90%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ArgSer 182
                                                                                                                                                                                                                                                                                                                                                                                                Percent Similarity:
Best Local Similarity:
                                                                                                                                                                                                                                                                                                                                                      Alignment Scores:
Pred. No.:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    361
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  141
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              421
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           161
                                                                                                                                                                                                                                                                                                                                                                                                                             Query Match:
   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     QQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   DP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     gg
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ογ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ó
```

AGGICC 546

a

```
Score:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Q
D
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ογ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       à
                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                               ô
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 đ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
8288888
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                It is a double-stranded nucleotide sequence of the open reading frame (ORF) (tag a) extending through clones 14; 11b, 7f, 7e, 8h 33c, 40b, 35b, 36, 31b, 32c, 14c, 8f, 31f, 31g, and 39c of hepatitis C virus (HCV) CDNA. In creating the composite sequence the following heterogenelities were considered. Clone 33c contains a sequence of 800 base pairs which overlaps the cONAs in clones 40b and 37c. In clone 33c, as well as in 5 other overlapping clones, nucleotide #789 is a G. However, in clone 37b the corresponding nucleotide is an A. This heterogenity may have important ramifications for protein folding. Nucleotide #2 in clone 8h is a T which may represent a cloning artifact because the corresponding residue in clone 7e and in 3 other overlapping clones is an A. Therefore the residue in this position is designated as a C. Then a G because the corresponding residue in clone 31f is represented as a T than a G because the corresponding residue in clone 31f and in 2 other overlapping clones is a T. The 3' terminal sequence of clone 33f is represented as ATTC, as is found in the corresponding sequence in clone 31g and in 2 other overlapping clones 33f, no clone 33f. Residue #4 in clone 33f and 2 other corresponding residue in clone 33f
                                                                                                                                                                                                 Combined open reading frames of the hepatitis C virus (HCV) CDNA in clones 14i 11b, 7f, 7e, 8h, 33c, 40b, 37b, 35, 36, 81, 32, 33b, 25c, 14c,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                closes is an A. The 3'-terminus of close 141 is depicted as TA rather than AA because the corresponding dinucleotide in clone 11b and 3 other closes is TA. Potential cloning artifacts have been omitted and instead the corresponding sequences in non-5'-terminal regions of multiple overlapping clones are shown. AANO2097 could be used as a source of oligomeric DNA hybridisation probes to detect the presence of HCV nucleic acids in samples. The polypeptide(s) it encodes could be used as
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Claim 3; Figure 26-1, 26-2, 26-3, 26-4, 26-5, 26-6; 139pp; English.
                                                                                                                                                                                                                                                                                Hepatitis C virus; HCV; non-A, non-B hepatitis; NANBH
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             and associated nucleic acids and polypeptide(s)
                                                                                                                                                                                                                                                                                                                                                             Location/Qualifiers
3..5300
/*tag= a
                                                         AAN92097 standard; DNA; 5300 BP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Kuo G;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      87US-0139886.
88US-0161072.
88US-0191263.
88US-0263584.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     87US-0122714
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              8BEP-0310922
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        88US-0271450
                                                                                                                                                            (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Purified hepatitis C virus
                                                                                                                                           (nbdated)
                                                                                                                                                                                                                    clones 14 11b, 7f, 768f, 33f, 33g and 39c.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Houghton M, Choo QL,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              WPI; 1989-159274/22.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (CHIR ) CHIRON CORP
                                                                                                                                                                                                                                                                                                                           Hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   P-PSDB; AAP92041
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   18-NOV-1987;
30-DEC-1987;
26-FEB-1988;
06-MAY-1988;
26-OCT-1988;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              18-NOV-1988;
                                                                                                                                         25-MAR-2003
02-MAR-1990
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        31-MAY-1989
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        14-NOV-1988
                                                                                                                                                                                                                                                                                                                                                                                                                                               EP318216-A
                                                                                                    AAN92097;
```

```
1229
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1289
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1290 GGTGATAGCAGGGCAGCCTGCTGTCGCCCCGGCCCATTTCCTACTTGAAAGGCTCCTCG 1349
                                                                                                                                                                                                                                                                                                        1049
                                                                                                                                                                                                                                                                                                                                                                                   160
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       120
                                                                                                                                                                                                                                                                                                                                                                                                                                      81 LysAspLeuValGlyTrpProAlaProGlnGlySerArgSerLeuThrProCysThrCys 100
                                                                                                                                                                                                                                                                                40
                                                                                                                                                                                                                                                                                                                                  9
immuno- assay reagents and vaccines and to generate antibodies useful in diagnosis and passive immunotherapy for HCV infection/non-A, non-B
                                                                                                                                                                                                                                                                               121 GlyAspSerArgGlySerLeuLeuSerProArgProlleSerTyrLeuLySGlySerSer
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          141 GlyGlyProLeuLeuCysProAlaGlyHisAlaValGlyIlePheArgAlaAlaValCys
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        101 GlySerSerAspLeuTyrLeuValThrArgHisAlaAspValIleProValArgArgArg
                                                                          Sequence 5300 BP; 1047 A; 1606 C; 1515 G; 1130 T; 2 other;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Hepatitis C virus; composite cDNA; probe; vaccine.
                                                                                                                Length:
Matches:
Conservative:
Mismatches:
                                     (Updated on 25-MAR-2003 to correct PR field. (Updated on 25-MAR-2003 to correct PI field.
                                                                                                                                                                 Indels:
                                                                                                                                                                               Gaps:
                                                                                                                                                                                                      US-09-965-594-1 (1-182) x AAN92097 (1-5300)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Hepatitis C virus composite probe.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          AAN90327 standard; cDNA; 5360 BP
                                                                                                                8.9e-78
943.00
100.00%
98.35%
98.95%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (npdated)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1470 AGGTCC 1475
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             181 ArgSer 182
                                                                                                                                         Percent Similarity:
Best Local Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Pan troglodytes
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           25-MAR-2003
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       11-NOV-1989
                                                                                                     Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    AAN90327;
                                                                                                                                                                   Query Match:
                                                                                                                  ..
Q
```

```
It is a double-stranded nucleotide sequence of the open reading frame (ORF) (tag a) extending through clones 12f to 15e of hepatitis C virus (HCV) cDNA. It can be used to make oligomeric DNA hybridlsation probes t detect the presence of HCV nucleic acids in samples. The polypeptide(s) it encodes could be used as immunoassay reagents and vaccines and to generate antibodies useful in diagnosis and passive immunotherapy for HCV infection/non-A, non-B hepatitis. (Updated on 25-MAR-2003 to correct PR field.)
                                                                                                                                                                                                                                                                                                                                            Combined open reading frames of the hepatitis C virus (HCV) cDNAs from
          141 GlyGlyProLeuLeuCysProAlaGlyHisAlaValGlyIlePheArgAlaAlaValCys
                                                                 GlyAspSerArgGlySerLeuLeuSerProArgProIleSerTyrLeuLysGlySerSer
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Sequence 6905 BP; 1421 A; 2082 C; 1946 G; 1456 T; 0 other;
                                                                                                                                                                                                                                                                                                                                                                                     non-B hepatitis; NANBH
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     and associated nucleic acids and polypeptide(s)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Claim 3; Figure 32-1 - 32-7; 139pp; English
                                                                                                                                                                                                                                                                                                                                                                                                                                       Location/Qualifiers
3..6905
/*tag= a
                                                                                                                                                                                                                                                BP.
                                                                                                                                                                                                                                                                                                                                                                                     Hepatitis C virus; HCV; non-A,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ö
                                                                                                                                                                                                                                                standard; DNA; 6905
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           87US-0122714.
87US-0139886.
88US-0161072.
88US-0263584.
88US-0263584.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   88EP-0310922
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Kuo
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Purified hepatitis C virus
                                                                                                                                                                                                                                                                                                      (updated)
(first entry)
                                                                                                                                                                                                                                                                                                                                                            clones 12f through 15e.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Choo QL,
                                                                                                                                                                                  1470 AGGICC 1475
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                WPI; 1989-159274/22.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          (CHIR ) CHIRON CORP
                                                                                                                                                                ArgSer 182
                                                                                                                                                                                                                                                                                                                                                                                                               Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               P-PSDB; AAP92047
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             18-NOV-1987;
30-DEC-1987;
26-FEB-1988;
06-MAY-1988;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    26-OCT-1988;
14-NOV-1988;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Houghton M,
                                                                                                                                                                                                                                                                                                     25-MAR-2003
02-MAR-1990
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   18-NOV-1988;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        31-MAY-1989
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              EP318216-A.
                        1290
                                                                              1350
                                                                                                                                                                181
                                                                                                                                                                                                                                                                           AAN92103;
121
                                                                                                                                                                                                                                               AAN92103
                                                                                                                                                                                                                                                                                                                                                                                                                                          Key
                                                                                                                                                                                                                      RESULT 7
                                                                                                                                                                                                                                     AAN92103
                                                                                                                                                                                                                                                             d
                                                   ò
                                                                             a
                                                                                                           ò
                                                                                                                                                                ö
                       a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1109
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1169
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1229
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       100
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        GlySerSerAspLeuTyrLeuValThrArgHisAlaAspValIleProValArgArgArg 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      989
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            08
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      96
                                                                                                                                                                                                                                                                                                                                                                     The sequence shows the composite cDNA sequence derived from the aligned hepatitis C virus (HCV) cDNA's in clones 14i, 11b, 7f. 7e, 8h, 33c, 40b, 37b, 35, 36, 81, 32, 33b, 25c, 14c, 8f, 33f, 33f and 39c. The cDNA encodes antigens which react with antibodies in patients with non-A non-B hepatitis (NANBH). The cDNA can be used to design probes, or to synthesise polypeptides, which are used to dispasse HCV; induced NANBH, to raise antibodies for immunoassay or treatment, or to produce vaccines, see also AAP90158, AAN90303-26, and AAN90328-36.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ClyThrArgThrIleAlaSerProLysGlyProVallieGInMc.TyrThrAsnValAsp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1110 GGAACGAGGACCATGGCGTCACCCAAGGGTCCTGTCATCCAGATGTATACCAATGTAGAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Lysaspi.euValGiyTrpProAlaProGinGlySerArgSer..euThrProCysThrCys
                                                                                                                                                                                                                                                                               Hepatitis C virus gene - used for prodn. of polynucleotide probes, polypeptide(s) and antibodies for diagnosis, prevention and treatment of infection.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Sequence 5360 BP; 1060 A; 1622 C; 1532 G; 1145 T; 1 other;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   5360
179
3
0
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                (Updated on 25-MAR-2003 to correct PR field.)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Length:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Gaps:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  US-09-965-594-1 (1-182) x AAN90327 (1-5360)
                                                                                                                                                                                                                                                                                                                                              Disclosure; Fig. 26; 174pp; English.
Location/Qualifiers
                                                                                                                                                                                                                                   Ö
                                                                                                                                   8705-0122714.
8705-0139886.
8805-0161072.
8805-0263584.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 9.03e-78
943.00
100.00%
98.35%
98.95%
                                                                                                           88GB-0027024
                                                                                                                                                                                                                                     Kuo
                                                                                                                                                                                                         (CHIR ) CHIRON CORPORATION
                           ø
          3..5360
/*tag-
                                                                                                                                                                                                                                  Choo QL,
                                                                                                                                                                                                                                                            WPI; 1989-215054/30.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Percent Similarity:
Best Local Similarity:
                                                                                                                                                30-DEC-1987;
26-FEB-1988;
26-OCT-1988;
                                                                                                                                                                                                                                  Houghton M,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Alignment Scores:
                                                                                                           18-NOV-1988;
                                                    GB2212511-A.
                                                                                26-JUL-1989
                                                                                                                                      18-NOV-1987
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      930
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  81
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        101
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Query Match:
DB:
Key
```

ŏ g ò a ò a ò a ò Q ò

ដ

Kuo G;

```
It is a double-stranded nucleotide sequence of the open reading frame (ORF) (tag a) extending through clones K9-1 to 15e of hepatitis C virus (HCV) cobh. It can be used to make oligomeric DNA hybridisation probes to detect the presence of HCV nucleic acids in samples. The polypeptide(s) it encodes could be used as immunoassay reagents and vaccines and to generate antibodies useful in diagnosis and passive immunotherapy for HCV infection/non-A, non-B hepatitis.
(Updated on 25-MAR-2003 to correct PR field.)
                                                                                                                                                                                                                                                                                                                       and associated nucleic acids and polypeptide(s)
                                                                                                                                                                                                                                                                                                                                                       Claim 3; Figure 47-1 - 47-8; 139pp; English
                                               870S-0122714.
870S-0139886.
860S-0161072.
880S-0191263.
88US-0263584.
88US-0271450.
               88EP-0310922
                                                                                                                                                                                                                                                                                                    Purified hepatitis C virus
                                                                                                                                                                                                           Houghton M, Choo QL,
                                                                                                                                                                                                                                               WPI; 1989-159274/22
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Best Local Similarity:
                                                                                                                                                                                                                                                                  P-PSDB; AAP92050
                                                                                                                                                                           CHIR ) CHIRON
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Percent Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Alignment Scores:
                                               18-NOV-1987;
30-DEC-1987;
                                                                                                                                       14-NOV-1988;
               18-NOV-1988;
                                                                                   26-FEB-1988
06-MAY-1988
                                                                                                                        26-OCT-1988
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              21
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1908
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1968
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       2028
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               2088
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     101
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Query Match:
DB:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ٠.
و
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Dp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         òγ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           a
1562
                                                                                                                                                                                                                                                                                 1322
                                                                                                                                                                                                                                                                                                                                                                                                                                                               1683 ACCCGTGGAGTGGCTAAGGCGGTGGACTTTATCCCTGTGGAGACCTAGAGACAACCATG 1742
                                                                                                                                                                       101 GlySerScraspLeuTyrLeuValThrargHisAlaaspvallieProValargArgarg 120 [110] [111] [111] [111] [111] [111] [111] [111] [111] [111] [111] [111] [111] [111] [111] [111] [111] [111] [111] [111] [111] [111] [111] [111] [111] [111] [111] [111] [111] [111] [111] [111] [111] [111] [111] [111] [111] [111] [111] [111] [111] [111] [111] [111] [111] [111] [111] [111] [111] [111] [111] [111] [111] [111] [111] [111] [111] [111] [111] [111] [111] [111] [111] [111] [111] [111] [111] [111] [111] [111] [111] [111] [111] [111] [111] [111] [111] [111] [111] [111] [111] [111] [111] [111] [111] [111] [111] [111] [111] [111] [111] [111] [111] [111] [111] [111] [111] [111] [111] [111] [111] [111] [111] [111] [111] [111] [111] [111] [111] [111] [111] [111] [111] [111] [111] [111] [111] [111] [111] [111] [111] [111] [111] [111] [111] [111] [111] [111] [111] [111] [111] [111] [111] [111] [111] [111] [111] [11] [111] [111] [111] [111] [111] [111] [111] [111] [111] [111] [111] [111] [111] [111] [111] [111] [111] [111] [111] [111] [111] [111] [111] [111] [111] [111] [111] [111] [111] [111] [111] [11] [111] [111] [111] [111] [111] [111] [111] [111] [111] [111] [11] [111] [11] [11] [11] [11] [11] [11] [11] [11] [11] [11] [11] [11] [11] [11] [11] [11] [11] [11] [11] [11] [11] [11] [11] [11] [11] [11] [11] [11] [11] [11] [11] [11] [11] [11] [11] [11] [11] [11] [11] [11] [11] [11] [11] [11] [11] [11] [11] [11] [11] [11] [11] [11] [11] [11] [11] [11] [11] [11] [11] [11] [11] [11] [11] [11] [11] [11] [11] [11] [11] [11] [11] [11] [11] [11] [11] [11] [11] [11] [11] [11] [11] [11] [11] [11] [11] [11] [11] [11] [11] [11] [11] [11] [11] [11] [11] [11] [11] [11] [11] [11] [11] [11] [11] [11] [11] [11] [11] [11] [11] [11] [11] [11] [11] [11] [11] [11] [11] [11] [11] [11] [11] [11] [11] [11] [11] [11] [11] [11] [11] [11] [11] [11] [11] [11] [11] [11] [11] [11] [11] [11] [11] [11] [11] [11] [11] [11] [11] [11] [11] [11] [11] [11] [11] [11] [11] [11] [11] [11] [11] [11] [11] [11] [11] [11] [11] [11] [11] [11] [11] [11] [11] [11] [11] [11] [11] [11] [11] [11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          140
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              141 GlyGlyProLeuLeuCysProAlaGlyHisAlaValGlyIlePheArgAlaAlaValCys 160
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      161 ThrargGlyValalaLysAlaValaSpPheLleProValGLuSerLeuGluThrThrwet 180
                                                                                                                                                                                                                                               40
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Combined open reading frames of the hepatitis C virus (HCV) cDNAs from clones K9-1 through 15e.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            121 GlyAspSerArgGlySerLeuLeuSerProArgProIleSerTyrLeuLysGlySerSer
                                                                                                                                                                                                                                                                    21 SerbeuThrGlyArgAspLysAsnGlnValGluGlyGluValGlnIleValSerThrAla
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 hepatitis; NANBH
             6905
179
3
0
0
           Length:
Matches:
Conservative:
Mismatches:
Indels:
Gaps:
                                                                                                                                       US-09-965-594-1 (1-182) x AAN92103 (1-6905)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   non-B
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Location/Oualifiers 3..7310
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   HCV; non-A,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    AAN92106 standard: DNA; 7310
               23e-77
                                943.00
100.00%
98.35%
98.95%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (updated)
(first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1743 AGGTCC 1748
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ArgSer 182
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Hepatitis C virus:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Hepatitis C virus.
                                             Percent Similarity:
Best Local Similarity:
Query Match:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         25-MAR-2003
02-MAR-1990
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   31-MAY-1989
Alignment Scores
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               EP318216-A
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       AAN92106;
                ж
..
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Key
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   RESULT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          a
                                                                                                                                                                                                                                                                                                                                                                                                                         qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    οy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              δ
                                                                                                                                                                                                         a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ó
                                                                                                                                                                           ò
                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                   οy
                                                                                                                                                                                                                                                                                                                                               a
                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
```

```
1847
                                                                                                                                                                                                                                               1907
                                                                                                                                                                                                                                                                                                                                       2027
                                                                                                                                                                                                                                                                                                                                                                                     2087
                                                                                                                                                                                                                                                                                                                                                                                                                                  2147
                                                                                                                                                                                                                                                                                           1967
                                                                                                                                                        1787
                                                                                                                                                                                                                                                                                                                  100
                                                                                                                                                                                                                                                                                                                                                              120
                                                                                                                                                                                                                                                                                                                                                                                                                                                         141 GlyGlyProLeuLeuCysProAlaGlyHisAlaValGly1lePheArgAlaAlaValCys 160
                                                                                                                                                                               0
                                                                                                                                                                                                                           9
                                                                                                                                                                                                                                                                       80
                                                                                                                                                                                                                                                                                                                                                              1788 AGCCTAACTGGCCGGGACAAAAACCAAGTGGAGGTGAGGTCCAGATTGTGTGTAACTGT
                                                                                                                                                                                                                                                                       61 GlyThrArgThrIleAlaSerProLysGlyProValIleGlnMetTyrThrAsnValAsp
                                                                                                                                                                                                                                                                                                                   81 LysAspLeuValGlyTrpProAlaProGlnGlySerArgSerLeuThrProCysThrCys
                                                                                                                                                                               SerLeuThrGlyArgAspLysAsnGlnValGluGlyGluValGlnIleValSerThrAla
                                                                                                                                                                                                                           41 AlaGlnThrPheLeuAlaThrCysIleAsnGlyValCysTrpThrValTyrHisGlyAla
C; 2058 G; 1540 T; 4 other;
                               7310
179
3
0
0
                                Length:
Matches:
Conservative:
Mismatches:
                                                                            Indels:
                                                                                        Gaps:
                                                                                                             US-09-965-594-1 (1-182) x AAN92106 (1-7310)
Sequence 7310 BP; 1491 A; 2217
                             1.32e-77
943.00
100.00%
98.35%
98.95%
```

q

õ qq ò

20

```
141 GlyGlyProLeuLeuCysProAlaGlyHisAlaValGlyIlePheArgAlaAlaValCys 160
                                                                                                                                                                                                                                                          MetalaProlleThrAlaTyrAlaGlnGlnThrArgGlyLeuLeuGlyCysIleIleThr
                                                                                                                                                      GlyThrArgThrIleAlaSerProLysGlyProValIleGlnMetTyrThrAsnValAsp
                                                                                                                                                                                                                                 1968 CAAGACCTTGTGGGGTGGCCCGCTCCGCAAGGTAGCCGCTCATTGACACCCTGCACTTGC
                                                                                                                                                                                                                                                                                                               121 GlyAspSerArgGlySerLeuLeuSerProArgProIleSerTyrLeuLysGlySerSer
                                                                                                                                                                                                                                                                                                                                                                              41 AlaGlnThrPheLeuAlaThrCysIleAsnGlyValCysTrpThrValTyrHisGlyAla
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Hepatitis C virus; HCV; antigen; detection; diagnosis; vaccine; antibodies; immunoprophylaxis; sera; serum; ds.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                BP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                standard; cDNA to mRNA; 7310
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Hepatitis C virus clone genome.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Reyes GR;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  91US-0681703.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             900S-0505611.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  (updated)
(first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (GENE-) GENELABS INC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              2268 AGGTCC 2273
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Moeckli R,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      WPI; 1995-302120/39
                                                                                                                                                                                                                                                                                                                                                                                                                                                                     181 ArgSer 182
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  05-APR-1991;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             06-APR-1990;
09-OCT-1990;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   25-MAR-2003
15-AUG-1996
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 US5443965-A
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          22-AUG-1995
                                                                                                                                                                                                                                                                                                                                                                                         2148
-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         AAQ98221;
                                                                                                                                                       61
                                                                                                                                                                                                         81
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              AAQ98221
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Kim JP,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        RESULT 10
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       AAQ98221
                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                     δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             dC
                                                                                                                                                                                                                                                                                qq
                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                     a
                                                                                                                                                                                                                                                                                                                                                               ò
                                                                         qq
                                                                                                                       a
                                                                                                                                                       ò
                                                                                                                                                                              g
                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                            ò
                   8
                                              ΟŽ
                                                                                                   ö
                                                  ď
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             The sequence shows a composite hepatitis C virus (HCV) cDNA, derived be aligning clones k9-1 through 15e in 5'-3' direction. The CDNA encodes antiquens which react with antibodies in patients with non-a non-B hepatitis (NANBH). The CDNA can be used to design probes, or to synthesise polypeptides, which are used to diagnose HCV-induced NANBH.
Hepatitis C virus gene - used for prodn. of polynucleotide probes, polypeptide(s) and antibodies for diagnosis, prevention and treatment of infection.
                                     161 ThrArgGlyValAlaLysAlaValAspPhcIleProValGluSetLeuGluThrThrMet
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               to raise antibodies for immunoassay or treatment, or to produce vaccines. See also AAP90288, and AAM90303-35.

(N.B. This record was resubmitted to correct crrors in the sequence.) (Updated on 25-MAR-2003 to correct PR field.)
                                                                                                                                                                                                                                                                                                 Hepatitis C virus; cDNA; clone 15e; clone k9-1; probe; vaccine; ds.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Sequence 7310 BP; 1495 A; 2218 C; 2058 G; 1539 T; 0 other;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    7310
179
3
0
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Length:
Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Gaps:
                                                                                                                                                                                                                                                                         Composite hepatitis C virus (HCV) CDNA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Disclosure; fig 47; 235pp; English.
                                                                                                                                                                   AAN90336 standard; DNA; 7310 BP
                                                                                                                                                                                                                                                                                                                                                                                                                             87US-0122714.
87US-0139886.
88US-0161072.
88US-0263584.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Kuo G;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                943.00
100.00%
98.35%
98.95%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1.32e-77
                                                                                                                                                                                                                     (updated)
(updated)
(first entry)
                                                                                                                                                                                                                                                                                                                                                                                                        88GB-0027024
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               (CHIR ) CHIRON CORPORATION.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Houghton M, Choo QL,
                                                                                                                AGGTCC 2273
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               WPI; 1989-215054/30.
P-PSDB; AAP90288.
                                                                                        181 ArgSer 182
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Percent Similarity:
Best Local Similarity:
Query Match:
                                                                                                                                                                                                                                                                                                                           Pan troglodytes
                                                                                                                                                                                                                                                                                                                                                                                                                                           30-DEC-1987;
26-FEB-1988;
26-OCT-1988;
                                                                                                                                                                                                                                                                                                                                                                                                       18-NOV-1988;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Alignment Scores:
                                                                                                                                                                                                                                 19-JUL-2001
01-NOV-1989
                                                                                                                                                                                                                      25-MAR-2003
                                                                                                                                                                                                                                                                                                                                                    3B2212511-A.
                                                                                                                                                                                                                                                                                                                                                                             26-JUL-1989
                                                                                                                                                                                                                                                                                                                                                                                                                                18-NOV-1987
                                                                                                                  2268
                                                                                                                                                                                              4AN90336;
```

AAN90336

XXX
AAN9
DT 19-J
DT

2087

140

2027

100

80

2207

to of

New nucleic acids encoding hepatitis C virus antigens - used develop prods. for detection of HCV-infected sera and prodn.

(1-7310)

US-09-965-594-1 (1-182) x AAN90336

.. №

Score:

```
CDNA sequence compiled Hepatitis C virus cDNA clones
                                                                                                                                                                                                                                                                                                                                                                                                                                                   Alignment Scores:
                                                                                                      EP1034785-A2
                                                                                                                                                               20-APR-1989;
18-MAY-1989;
                                                                                                                                                                                 16-MAR-1990;
                                                                                                                                                        17-MAR-1989;
                                                                                                                      13-SEP-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        -
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Query Match:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Score:
 윱
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ô
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      2087
                                                                                                                                                                                                                                                                                                                                                                                                                                                                    2027
                                                                                                                                                                                                                                                                                                                                                                                                                                          2088 GGTGATAGCAGGGGCAGCCTGCTGTGGCCCCGGCCCATTTCCTACTTGAAAGGCTCCTCG 2147
                                                                                                                                                                                                                                                                                                                                                                                                                         760
                                                                                                                                                                                                                                                                                                                                                                                        101 GlySerSerAspLeuTyrLeuValThrArgHisAlaAspVallleProValArgArgArg 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             161 ThrArgGlyValAlaLysAlaValAspPheIJeProValGluSerLeuGluThrThrMet 180
                                                                                                                                                                                                                                                                                                                                                       81 LysaspLeuValGlyTrpProAlaProGlnGlySerArgSerLeuThrProCysThrCys 100
                                                                                                                                                                                                                                                    40
                                                                                                                                                                                                                                                                                    09
                                                                                                                                                                                                                                                                                                                                                                 141 GlyGlyProLeuLeuCysProAlaGlyHisAlaValGly11ePhcArgAlaAlaValCys
                                                                                                                                                                                                                                                     Sert.euThrGlyArgAspi.ysAsnGlnValGluClyGluValGlnIleValSerThrAla
                                                                                                                                                                                                                                                                                    41 AlaGlnThrPheLeuAlaThrCysIleAsnGlyValCysTrpThrValTyrHisGlyAla
                                Hepatilis C virus (HCV) antigens can be used for detecting HCV infected sera and individuals infected with HCV. They can also be used in an anti-HCV vaccine or for the production of anti-HCV antibodies which can be used for passive immunoprophylaxis. The antigens consistently identify more HCV positive serum samples with a high degree of specificity. See AAQ098202-14 and AAR81939-51. (Updated on 25-MAR-2003 to correct PF field.)
                                                                                                              Sequence 7310 BP; 1494 A; 2217 C; 2060 G; 1539 T; 0 other;
                                                                                                                                      7310
179
3
0
0
                                                                                                                                      Length:
Matchcs:
Conscrvative:
Mismatches:
Indels:
Gaps:
                                                                                                                                                                                                  US-09-965-594-1 (1-182) x AAQ98221 (1-7310)
                   Example 4; Figure 11; 71pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ВР
 vaccines and anti-HCV antibodies
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CDNA; 8316
                                                                                                                                      1.32e-77
943.00
100.00%
98.35%
98.95%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   15-JAN-2001 (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         2268 AGGTCC 2273
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               181 ArgSer 182
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  standard;
                                                                                                                                                        Percent Similarity:
Best Local Similarity:
                                                                                                                               Alignment Scores:
Pred. No.:
                                                                                                                                                                                                                                   1728
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  AAA75296
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    AAA75296;
                                                                                                                                                                                                                                                    21
                                                                                                                                                                          Query Match:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         AAA75296

1D AAA7

XX

AC AAA7

XX

DT 15-3
                                                                                                                                                                                                                                                                                                                                                                                                                                         qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Dp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                       Ω
                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
 ò
                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                  a
                                                                                                                                                                                                                                                                                                                   ŏ
                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                      QQ
                                                                                                                                                                                                                                                                                                                                                                                         ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                            à
```

```
2853
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         2734 CIGGCGCCCAICACGCCTACGCCCAGCAGACAAGGGGCCICCTAGGGTGCATAAICACC 2793
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           40
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     The specification describes a pharmaceutical composition which comprises a hepatitis C virus (HCV) ancisence polynociacided. The HCV is characterized by a positive stranded RNA genome which has 40% homology at the polypeptide level to a HCV polyproctain. The antisense polynociacide binds to cellular polynociacidetides which enhance and/or are required for viral infectivity, replicative ability or chronicity. The antisense polynocleotides may also be designed to bind with high specificity, to be of increased stability, to be stable and to have low toxicity. The composition also comprises an agent which causes viral RNA to be inactive. The composition is used for preventing HCV replication in a system. The present course of the invention.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     MetAlaProlleThrAlaTyrAlaGlnGlnThrArgGlyLeuLeuGlyCysIleIleThr
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              41 AlaGlnThrPheLeuAlaThrCysIleAsnGlyValCysTrpThrValTyrHisGlyAla
                                                                                                                                                                              /*tag= a
/note= "partial sequence; no termination codon given"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Novel composition comprising a hepatitis C virus antisense polynucleotide which is complementary to or corresponds to a sense strand of the virus genome, and selectively hybridises to it -
Hepatitis C virus; HCV; antisense polynucleotide; polyprotein;
viral infectivity; viral replication; ds.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Sequence 8316 BP; 1671 A; 2529 C; 2345 G; 1771 T; 0 other;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     8316
179
3
0
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Matches:
Conservative:
Mismatches:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Length:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       US-09-965-594-1 (1-182) x AAA75296 (1-8316)
                                                                                                                           Location/Qualifiers
1..8316
/*tag= a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Example; Fig 16; 75pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                      89US-0325338.
89US-0341334.
89US-0355002.
90EP-0302866.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1.55e-77
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             943.00
100.00%
98.35%
98.95%
                                                                                                                                                                                                                                                                                                                                                           16-MAR-1990; 2000EP-0109602
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Kuo
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Houghton M, Choo Q,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     WPI; 2000-566891/53.
P-PSDB; AAB18540.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       (CHIR ) CHIRON CORP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Percent Similarity:
Best Local Similarity:
                                                                              Hepatitis C virus.
```

us-09-965-594-1.rng

```
Score
                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                               q
                                                                                                                                                                                                                                                                                                                                                                       q
                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                6
                                                                                                                                                                                                                                                                                                                                                                                                             õ
                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                       gg
                                                                                                                                                                                                                                                                                             q
                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                      q
                                                                                                                                                                                                                                                                                                                                                             ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                      ó
                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
ò
                                                                                                                                                                                                                                   ^{\circ}
                                                                                                                                                                                                                                              q
                                                                                                                                                                                                                                                                                   ô
                                                                                                    /transl_except. (pos:1588..1589; aa:Leu)
/note= 'this codon has an apparent 1 nucleotide deletion,
which alters the reading frame'
/transl_except= (pos:1647..1650; aa:Pro)
/note= 'this codon has an apparent 1 nucleotide
insertion, which alters the reading frame: insertion, which alters the reading frame: present in the formal sequence
present in the formal sequence listing of the
specification.
                                                      101 GlyScrSerAspLeuTyrLeuvalThrArgHisAlaAspValIleProValArgArg 120
                                                                                                                                                                                                                                               Hepatitis C virus; HCV; J1; J7; HCV-1: non-A, non-B HCV; NANBH;
HCV infection; vaccine; ds.
                                                                                                                                                                                                                                                                                                                                                                                                                                            Kolberg JA:
                                                                                                                                                                                                                                                                                                                                                                                                                                            Irvine BD,
                                                                                                                                                                                                                                                                              Location/Qualifiers
268..9132
                                                                                                                                                                                                                                                                                                                                                                                                                                            Houghton M, Irvi
ito I, Weiner AJ;
                                                                                                                                                                                                                                    Nucleotide sequence of HCV-1 ORF.
                                                                                                                                                                                          AAZ07656 standard; DNA; 9133 BP
                                                                                                                                                                                                                                                                                                                                                                                               89US-0408045.
89US-0456142.
90EP-0310149.
                                                                                                                                                                                                                                                                                                                                                                                     99FP-0101746
                                                                                                                                                                                                                        (first entry)
                                                                                                                                                                                                                                                                                           /*tag=
                                                                                                                                                                                                                   (updated)
                                                                                                                                                                                                                                                                                                                                                                                                                                       Han J, n. Saito J
                                                                                                                                                                 3274 AGGTCC 3279
                                                                                                                                                                                                                                                                                                                                                                                                                         (CHIR ) CHIRON CORP.
                                                                                                                                                     181 ArqSer 182
                                                                                                                                                                                                                                                                  Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                               15-SEP-1989;
21-DEC-1989;
17-SEP-1990;
                                                                                                                                                                                                                                                                                                                                                                                    17-SEP-1990;
                                                                                                                                                                                                                                                                                                                                                                                                                                                  Miyamura T,
                                                                                                                                                                                                                 20-MAR-2003
08-NOV-1999
                                                                                                                                                                                                                                                                                                                                                             EP939128-A2
                                                                                                                                                                                                                                                                                                                                                                        01-SEP-1999
                                                                                                                                                                                                      AAZ07656;
                                                                                                                                                                                                                                                                                                                                                                                                                                            Cha T,
                                                                                                                                                                                   q
                                                                                                  ò
                                                                                                               g
                                                                                                                         ò
                                                                                                                                      QQ
                                                                                                                                                    ò
                 a
                             ò
                                       g
                                                     ò
                                                               원
                                                                              ò
                                                                                       d
```

```
3582
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   3642
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         3702
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  3762
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             3462
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               140
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 100
                                                                                                                                                         and JT. These two isolates comprise nuclectide and anino acid sequences that are distinct from the HCV isolate HCV-1. The nucleotide sequences any be used to detect non-A, non-B HCV (NAMBH) polynucleotides by screen blood donors, donated blood and blood products for this infection. The isolates may also be used to isolate other naturally occurring variants of the virus. The polypeptides may be used as a vaccine for administration to patients to protect against infection with NAMBH. The Updated on 20-MAR-2003 to correct PF field.)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         80
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       40
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     41 AlaGlnThrPheLeuAlaThrCysIleAsnGlyValCysTrpThrValTyrHisGlyAla
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     3583 CARGACCTTGTGGCCCCCCCTCCCCAAGGTACCCCCTCATTGACCTCCACTTGC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1 MetalaProlleThralaTyralaGlnGlnThrargGlyLeuLeuGlyCysIlelleThr
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 81 LysaspLeuValGlyTrpProAlaProGlnGlySerArgSerLeuThrProCysThrCys
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       141 GlyGlyProLeuLeuCysProAlaGlyHisAlaValGlyIlePheArgAlaAlaValCys
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                161 ThrargGlyValAlaLysAlaValAspPheIleProValGluSerLeuGluThrThrMet
                                                                                                                                               the invention provides two new isolates of hepatitis C virus (HCV), J1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     3463 GCCCAAACCTTCCTGGCAACGTGCATCAATGGGGTGTGCTGCACTGTCTACCACGGGGCC
                                                      C Virus isolates, useful for diagnosis of hepatitis
                                                                                                                                                                                                                                                                                                                                                                                         Sequence 9133 BP; 1834 A; 2772 C; 2600 G; 1927 T; 0 other;
                                                                                                                                                                                                                                                                                                                                                                                                                                              91133
1179
0
0
                                                                                                                                                                                                                                                                                                                                                                                                                                            Length:
Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             US-09-965-594-1 (1-182) x AA207656 (1-9133)
                                                                          infections and development of vaccines
                                                                                                            Disclosure; Fig 12; 132pp; English
                                                                                                                                                                                                                                                                                                                                                                                                                                          1.74e-77
943.00
100.00%
98.35%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Percent Similarity:
Best Local Similarity:
Query Match:
DB:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             181 ArgSer 182
WPI; 1999-480843/41.
P-PSDB; AAY14975.
                                                        New Hepatitis
                                                                                                                                                                                                                                                                                                                                                                                                                            Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            3523
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           61
```

3883 AGGICC 3888

3514

40

3574

80

9

RESULT 13 AAQ05956

```
121 GlyAspSerArgGlySerLeuLeuSerProArgProIleSerTyrLeuLysGlySerSer 140
                                                                                      3395 CTGGCGCCCATCACGGCGTACGCCCAGCAGACAGGGGCCTCCTAGGGTGCATAATCACC 3454
                                                                                                                                  3755 GGTGATAGCAGGGAGCCTGCTGTCGCCCCGGCCCATTTCCTACTTGAAAGGCTCCTCG
                                                                                                                                                                            41 AlaGinThrPheLeuAlaThrCysIleAsnGlyValCysTrpThrValTyrHisGlyAla
                                                                                                                                                                                                                                     61 GlyThrArgThrIleAlaSerProLysGlyProValIleGlnMetTyrThrAsnValAsp
                                                        21 SerLeuThrGlyArgAspLysAsnGlnValGluGlyGluValGlnIleValSerThrAla
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Hepatitis C virus; HCV-1; non-A, non-B hepatitis; HCV antigen;
viral infections; 9s.
Gaps:
                            US-09-965-594-1 (1-182) x AAQ05956 (1-9185)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        AAQ10566 standard; DNA; 9185 BP.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Hepatitis C virus strain 1 DNA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             90EP-0309120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         890S-0398667
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  (updated)
(first entry)
1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              3935 AGGTCC 3940
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (CHIR ) CHIRON CORP.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                181 ArgSer 182
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 WPI; 1991-059670/09
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             21-AUG-1990;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           25-AUG-1989;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 25-MAR-2003
29-APR-1991
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 27-FEB-1991
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   3P414475-A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Weiner AJ,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     AAQ10566;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            RESULT 14
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          AAQ10566
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        qq
                                                                                                                                                                                                                                                                                                                                                                                                                                             q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                 a
                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                           ^{\circ}_{\lambda}
                                                                                                                                                                                                                                                                                                                                                                                                                   οy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                      a
                                                                                                                                             a
                                                                                                                                                                            ò
                                                                                                                                                                                                       q
                                                                                                                                                                                                                                     ζ
                                                          à
                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                HCV CDNA libraries were constructed using pooled serum from a chimpanzee with chronic HCV infection. A lambda 9111 library was screened with probes derived from previously isolated clones. The ORF is derived from the overlapping clones bilda, ag30a, CA205a, 12f, 14i, 11b, 7f, 8h, 33c, 40b, 37b, 35, 36, 81, 32, 33b, 25c, 14c, 8f, 31f, 39c, 35f, 19g, 26g, 15e, b5a and 16hr. These clones extend the sequence of the HCV genome reported in EP-118216. The upstream region from nucleotides -319 to -1348 (-1.1667 in this file) is covered by clones bilda, 18g, ag30a, CA205a, CA205a, CA205a, pida, cA167b, CA156e, CA84a and CA59a; nucleotides 8659-8866 (-8978-9185 in this file) are covered by clones b5a and
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Hepatitis C virus DNA - used for producing probes, polypeptide(s), antibodies and anti-sense polynucleotide(s) for diagnosis and therapy.
                                                                                                                                                                                                                                                                                                                within this region is claimed"
                                                                                                                                                  cDNA sequence.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Sequence 9185 BP; 1849 A; 2790 C; 2608 G; 1938 T; 0 other;
                                                                                                                                                                                                                                                                                                                                                         /note="encodes an epitope that is claimed"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Length:
Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                                                  Sense strand of the compiled Hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               See also AAQ05955.
(Updated on 25-MAR-2003 to correct PA field.)
                                                                                                                                                                              Hepatitis C virus (HCV); antiviral agent; ss.
                                                                                                                                                                                                                                       Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Disclosure, Fig 17; 83pp; English.
                                                                                                                                                                                                                                                                 /*tag~ a
1..1667
/*tag~ b
/note~*epitope w
8978..9185
                                          AAQ05956 standard; DNA; 9185 BP.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                            89US-0355002.
89US-0325338.
89US-0341334.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Kuo G;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1.75e-77
943.00
100.00%
98.35%
98.95%
                                                                                                                                                                                                                                                                                                                                                                                                                                                 90EP-0302866
                                                                                                       (updated)
(first entry)
                                                                                                                                                                                                                                                     320..9185
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Houghton M, Choo QL,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        (CHIR ) CHIRON CORP.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                WPI; 1990-284418/38
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Percent Similarity:
Best Local Similarity:
Query Match:
                                                                                                                                                                                                          virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                P-PSDB; AAR08124
                                                                                                                                                                                                                                                                                                                                                                                                                                                                              18-MAY-1989;
17-MAR-1989;
20-APR-1989;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                 16-MAR-1990;
                                                                                                      25-MAR-2003
23-JAN-1991
                                                                                                                                                                                                          Hepatitis C
                                                                                                                                                                                                                                                                                                                                                                                                                       19-SEP-1990
                                                                                                                                                                                                                                                                                                                                                                                          EP388232-A
                                                                         AAQ05956;
                                                                                                                                                                                                                                                                                    misc_RNA
                                                                                                                                                                                                                                                                                                                               misc_RNA
```

```
(first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                              Alignment Scores:
                                                                                                                EP1034785-A2
                                                                                                                                                                      20-APR-1989;
                                                                                                                                                                            18-MAY-1989;
16-MAR-1990;
       15-JAN-2001
                                                                                                                                13-SEP-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                      Pred. No.:
                                                                                                                                                                                                                                                                                                                                                                                                                                               Score:
q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         a
                                                                                                                                                                                                                                                                               3634
                                                                                                                                                                                                                                                                                                                                          3874
                                                                                                                                                                                                                                                                                                                                                                                                                                                                   120
                                                                                                                                                                                                                                                                                                                                                                                                        140
                                                                                                                                                                                                                                                                                                                                                                                                                                     160
                                                                                                                                                                                                                                                                                                             80
                                                                                                                                                                                                                  1 MetalaProlleThralaTyralaGinGinThrargGlyLeuLeuGlyCysIleTleThr 20
                                                                                                                                                                                                                                                                                                                      GlyThrargThr1lealaSerProLySGlyProVal1leGlnMetTyrThrAsnValAsp
                                                                                                                                                                                                                                                                                                                                                                        GlySerSerAspLeuTyrLeuValThrArgHisAlaAspValIleProValArgArg
                                                                                                                                                                                                                                                                                                                                                                                                        GlyAspScrArgGlySerLcuLeuSerProArgProlleScrTyrLeuLysGlySerSer
                                                                                                                                                                                                                                                                                                                                                                                                                                      GlyGlyProLeuLeuCysProAlaGlyHisAlaValGlyIlePheArgAlaAlaValCys
                                                                                                                                                                                                                                                                                                                                                                                                                                               This is a hepatitis C virus (HCV) composite chNA sequence, deduced using overlapping clones. a compsn. contg. the artigonic protein bodies (Abs) and for screening an agent which inhibits HCV aution. A cell line infected with this virus can be used as a action. A cell line infected with this virus can be used as a for treating viral infections. See also AAQ10567. (Updated on 25-MAR-2003 to correct PA field.)
     Cell lines infected with hepatitis C virus - are used as source of antigens for detection of HCV antibodies, for vaccines, and for screening anti-viral agents
                                                                                                                       Sequence 9185 BP; 1849 A; 2790 C; 2608 G; 1938 T; 0 other;
                                                                                                                                              9185
179
3
0
0
                                                                                                                                              Length:
Matches:
Conscrvative:
Mismatches:
Indels:
Gaps:
                                                                                                                                                                                                    US-09-965-594-1 (1-182) x AAQ10566 (1-9185)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   AAA75297 standard; cDNA; 9185 BP.
                                                                                                                                             1.75e-77
943.00
100.00%
98.35%
98.95%
                                     24pp;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  3935 AGGTCC 3940
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ArgSer 182
                                     fig l;
                                                                                                                                                                     Best Local Similarity:
Query Match:
DB:
                                                                                                                                                              Percent Similarity
                                                                                                                                       Alignment Scores:
                                                                                                                                                                                                                                                                                                                            3575
                                                                                                                                                                                                                                                                                                                                                                                                                                     141
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    181
                                                                                                                                                                                                                                                                                                             61
                                                                                                                                                                                                                                                                                                                                                                                                        121
                                                                                                                                                                                                                                  3395
                                                                                                                                                                                                                                                                                                                                                                          101
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  AAA75297
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   RESULT 15
AAA75297
                                                                                                                                               No.:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    a
ò
                                                                                                                                                                                                                                 å
                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                δ
                                                                                                                                                                                                                                                                                             qq
                                                                                                                                                                                                                                                                                                            ŏ
                                                                                                                                                                                                                                                                                                                          ដ្ឋ
                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                       a
                                                                                                                                                                                                                                                                                                                                                                        ô
                                                                                                                                                                                                                                                                                                                                                                                      q
                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                     à
                                                                                                                                                                                                                                                                                                                                                                                                                                                    q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                   õ
```

```
The specification describes a pharmaceutical composition which comprises a hepatitis C virus (HCV) antisense polynucleotide. The HCV is characterized by a positive stranded RNA genome which has antisense polynucleotide by a positive stranded RNA genome which has antisense polynucleotide binds to cellular polynucleotides which enhance and/or are required for viral infectivity, replicative ability or chronicity. The antisense polynucleotides may also be designed to bind with high specificity, to be of increased stability, to be stable and to have low toxicity. The composition also comprises an agent which coarses viral RNA to be inactive. The composition is used for preventing HCV replication in a system. The present sequence represents a novel HCV cDNA sequence, which is used in the
                                                                                                                                                                                                                          /*tag= a
/note= "partial sequence; no termination codon glven"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           sense
                                                HCV; antisense polynucleotide; polyprotein; viral replication; ds.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Novel composition comprising a hepatitis C virus antisense polynucleotide which is complementary to or corresponds to a strand of the virus genome, and selectively hybridises to it
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Sequence 9185 BP; 1849 A; 2790 C; 2608 G; 1938 T; 0 other;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      9185
179
.3
0
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Matches:
Conservative:
Mismatches:
Sense strand of HCV encoding a polyprotein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Length:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            US-09-965-594-1 (1-182) x AAA75297 (1-9185)
                                                                                                                                                                        Location/Qualifiers
320..9184
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Example; Fig 17; 75pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                             89US-0341334.
89US-0355002.
90EP-0302866.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ΰ
                                                                                                                                                                                                                                                                                                                                                                                                                                                        89US-0325338
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 943.00
100.00%
98.35%
98.95%
                                                                                                                                                                                                                                                                                                                                                                                                     16-MAR-1990; 2000EP-0109602
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Kuo
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   course of the invention.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Houghton M, Choo Q,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     WPI; 2000-566891/53.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (CHIR ) CHIRON CORP
                                                  Hepatitis C virus;
viral infectivity;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Best Local Similarity:
Query Match:
                                                                                                                          Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             P-PSDB; AAB18541
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Percent Similarity:
```

```
3935 AGGTCC 3940
                        181 ArgSer 182
     3575
   61
   6 6
                        õ
                          g
 Ω
õ
```

Search completed: August 30, 2003, 19:47:38 Job time: 188.009 secs

```
GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.
```

OM protein - protein search, using sw model

August 30, 2003, 18:01:52; Search time 9.01384 Seconds (without alignments) 949.524 Million cell updates/sec Run on:

US-09-965-594-1 953 | MADITAYAQOIRGLIGCLIT......GVAKAVDF.IPVESLETIMRS 182 Title: Perfect score: Sequence:

BLOSUM62 Gapop 10.0 , Gapext 0.5 Scoring table:

127863 seqs, 47026705 residues Searched:

127863 Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0 Maximum DB seq length: 2000000000

Post-processing: Minimum Match 6% Maximum Match 100% Listing first 45 summaries

SwissProt\_41:\* Database :

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

		æ			SUMMARIES	
Result		Query				
No.	Score		Length	DB	QI .	Description
	943	0.66	301	7	POLG_HCV1	4
7	927	97.3	3011	-	POLG_HCVH	7958 h genome
m	892		3010	~	POLG_HCVTW	5 h genome
4	891		3010	٦	POLG_HCVBK	3 h genome
ស	887		3010	~4	POLG_HCVJA	h genome
9	884		3010	٠4	POLG_HCVJT	h genome
7	714		3033	-	POLG_HCVJ8	h genome
ထ	712		3033	<b>~</b> 1	POLG_HCVJ6	h genome
6	87		209	- 4	PAAD_PSEAE	pscudomon
10	84		321	_	HHOA_ARATH	arabi
11	82		452	1	AAMP_HUMAN	
12			485	-4	Y136_TREPA	trepo
13	80.5		437	-	DEG1_ARATH	ara
14			253	~	CAC3_BOVIN	
15	4	7.8	415	-	ZP3_RABIT	ory
16	4		176	-	HYPF_AZOVI	
17	74.5		911		TB11_NEIMB	
18	1	7.8	326	-	PANE_RHILO	
19	73.5		263	-	GRAK_MOUSE	
20	~		730	-	HELS_METMA	Q8pzr7 methanosarc
21	72.5		257	-	GRAM_HUMAN	P51124 homo sapien
22	72	7.6	627	-	SAD1_MOUSE	mus 1
23	72		1527	. بـــ	CAIH_MOUSE	บรอน
24	72		2663	~	CENE_HUMAN	Q02224 homo sapien
25	72		3491	7	ERY1_SACER	Q03131 saccharopol
56	71.5		248	-	GRAD_MOUSE	
27	•	٠	323	-	VPRT_SMRVH	
28	71		219	_	SPR1_IPOBA	
53	71	7.5	336	~	UL16_EBV	P03221 epstein-bar
0,50	17	•	529	٦.	PGL2_RALSO	_
31	۲,	•	1180	٦,	ITAL_RAT	
32	70.5		264	п,	- 1	P40313 homo sapien
33	-	4./	659	7	VST2_HEVME	

P19532 homo sapien P06434 human cytom P06435 human cytom P28863 oryctolagus O8t139 methanosarc O91956 neisseria g P32434 schizosacch O8ye76 brucella me P05519 bacillus th P16480 bacillus th P16480 homo sapien P081160 homo sapien
TFE3_HUMAN VIE3_HCMVT VIE2_HCMVT MM03_RABIT HELS_METAC TBPL_METGO CWG2_SCHPO DNAK_BRUME C4BA_BACTI C4AA_BACTI PGBM_HUMAN CATG_HUMAN
нинненнанана
743 280 281 410 478 730 915 9355 637 1180 4391
70.5 70 70 70 70 69.5 69 69 69 69
88888444444 4889894444

## ALIGNMENTS

```
CELLUIAR ANINOPEPTIDASE.
CAPASID PROTEIN C (POTENTIAL).
MAJER PROTEIN (POTENTIAL).
MAJOR ENVELOPE PROTEIN E (POTENTIAL).
NONSTRUCTURAL PROTEIN NS.1/EZ (POTENTIAL).
PROTEASE/HELICASE NS3 (POTENTIAL).
PROTEASE/HELICASE NS3 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4A (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4A (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4B (POTENTIAL).
RNA DIRECTED RNA POLYMERASE (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                             SMART; SN00487; DEXDC; 1.

Polyprotein; Glycoprotein; Transferase; RNA-directed RNA polymerase;
Core protein; Coat protein; Envelope protein; Helicase; ATP-binding;
Transmembrane; Nonstructural protein; Hydrolase; Serine protease;
3D-structure. 1 | REMOVED FROM CAPSID PROTEIN C BY THE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                SIMILARITY).
SIMILARITY).
SIMILARITY).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       (POTENTIAL).
(POTENTIAL).
(POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        MW; 65F8C9447FCE5AF9 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              POTENTIAL.
CHARGE RELAY SYSTEM (BY
CHARGE RELAY SYSTEM (BY
CHARGE RELAY SYSTEM (BY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (GLCNAC.
(GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ATP (POTENTIAL).
DECH BOX.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            N-LINKED (N-LINKED (N-LINK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      N-LINKED
N-LINKED
N-LINKED
N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            N-LINKED
N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     N-L,INKED
                                                                                                                                                               Interpro; IPR001650; Helicase_C.
Interpro; IPR007095; RNA_pol_DS_PS.
Interpro; IPR007094; RNA_pol_PSvir.
                                                       HCV_NS3.
HCV_NS3.
HCV_NS4a.
HCV_NS4b.
HCV_NS5a.
                                                                                                                                                                                                                                                                                                                                  pF01006; HCV_NS4a; 1.
pF01001; HCV_NS4b; 1.
pF01506; HCV_NS5a; 1.
pF00271; helicase_C; 1.
pF00998; Viral_RdRP; 1.
om; pD186062; HCV_NS1; 1.
                                                                                                                                                                                                                  PPO1543 HCV_capsid) 1
PPO1542 HCV_capsid) 1
PPO1539 HCV_core: 1
PPO1559 HCV_MN: 1
PPO1558 HCV_MN: 1
PPO1538 HCV_MS: 1
PPO1907 HCV_MS: 1
PPO1006 HCV_MS: 1
PPO1001 HCV_MS: 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       327197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        645
2041
2077
2240
2364
2789
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1615
1862
2013
3011
                                                                      InterPro; IPR004109; H
InterPro; IPR000745; H
InterPro; IPR001490; H
InterPro; IPR002868; H
InterPro; IPR002166; H
             IPR002519;
IPR002531;
IPR002518;
IPR004109;
                                                                                                                                               InterPro; IPR002166;
InterPro; IPR001650;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Š
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            TRANSMEM
ACT_SITE
ACT_SITE
ACT_SITE
NP_BIND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                 Prodom;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CHAIN
CHAIN
CHAIN
CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CHAIN
                                                                                                                                                                                                                                                                                                                   Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                             Pfam;
                                                                                                                                                                                                                                                                                                                                                                            Pfam;
```

```
Ouery Match

99.0%; Score 943; DB 1; Length 3011;
Best Local Similarity 98.4%; Pred. No. 3.4e-82;
Matches 179; Conservative 3; Mismatches 0; Indels 0; Gaps 0;
```

```
GTRIIASPKGPVIQMYINVDKDLVGWPAPQGSRSLIPCTCGSSDLYLVIRHADVIPVRRR 120
                                                                                                                                             GDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVCTRGVAKAVDFIPVESLETTM 180
                                                                                                                                                                                                                                                                                                                                                                                                            01-AUG-1992 (Rel. 23, Created)
01-AUG-1992 (Rel. 23, Last sequence update)
15-SEP-2003 (Rel. 22, Last sequence update)
64-00-1992 (Rel. 23, Last sequence update)
65-00-1992 (Rel. 24, Last annotation update)
65-00-1992 (Rel. 25, Last annotation update)
65-00-1992 (GP18); Protein El (GP32) (GP35); Envelope glycoprotein E2
(GP68) (GP70) (NRI); Protein P7; Nonstructural protein NS2 (P21)
(EC 3.4.21.98); Nonstructural protein NS4A (P4): Nonstructural protein
NS4B (P27); Nonstructural protein NS5A (P56); Nonstructural protein
NS4B (P66) (P70) (RNA-directed RNA polymerase) (EC 2.7.7.48)].
Hepatitis C virus (isolate H) (HCV).
Viruses; SSRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              X-RAY CRYSTALLOGRAPHY (2.1 ANGSTROMS) OF 1207-1657.
MEDLINE-97331322; PubMed-9187654;
Yao N., Hesson T., Cable M., Hong Z., Kwong A.D., Le H.V., Weber P.C.;
"Structure of the hepatitis C virus RNA helicase domain.";
Nat. Struct. Biol, 4:463-467(1997).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         X-7AY CRYSTALLOGRAPHY (2.2 ANGSTROMS) OF 1192-1657.
MEDLINE-98154321; PubMed-9493270;
Kim J.L., Morgenstern K.A., Griffith J.P., Dwyer M.D., Thomson J.A.,
Murcko M.A., Lin C., Caron P.R.;
Hepatitis C virus NS3 RNA helicase domain with a bound
oligonucleotide: the crystal structure provides insights into the mode
of unwinding.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              -1- FUNCTION: PROTEASE WS2 IS RESPONSIBLE FOR THE CLEAVAGE OF NS2-NS3-1- FUNCTION: PROTEASE NS3 IS RESPONSIBLE FOR THE CLEAVAGE OF
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  -i- FUNCTION: NSSA SEEMS TO HAVE A TRANSCRIPTIONAL ACTIVATORY ROLE.
-i- FUNCTION: NSSB IS A RNA-DEPENDENT RNA POLYMERASE THAT FLAYS AN
ESSEMTIAL ROLE IN THE VIROS REPLICATION.
-i- CATALYTICA CALTUITY: Hydrolysis of four peptide bonds in the viral
precursor polyprotein, commonly with Asp or Glu in the P6
position, Cys or Thr in P1 and Ser or Ala in P1'.
-i- CATALYTIC ACTIVITY: N nucleoside triphosphate - N diphosphate +
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (RNA)(N).
SUBUNTT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS: E1 AND E2. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MRNA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        NS3-NS4A, NS4A-NS4B, NS4B-NS5A AND NS5A-NS5B.
-!- FUNCTION: NS4A FORMS A COMPLEX WITH NS3 AND IS ESSENTIAL FOR THE
ACTIVATION OF NS3.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   U
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               "Genomic structure of the human prototype strain H of hepatitis virus: comparison with American and Japanese isolates."; Proc. Natl. Acad. Sci. U.S.A. 88:10292-10296(1991).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SEQUENCE FROM N.A.
MEDLINE-92052256; PubMed-1658800;
Inchauspe G., Zebedee S., Lee D.H.H., Sugitani M., Nasoff M.
                                                                                                                                                                                                                                                                                                                                                                          PRT; 3011 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Structure 6:89-100(1998).
                                                                                                                                                                                                                                                                                                                                                                             STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         NCBI_TaxID-11108;
                                                                                                                                                                                                                                                                     RS 1207
                                                                                                                                                                                                                             181 RS 182
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Prince A.M.;
                                                                                                                                                                                                                                                                                                                                                                             POLG_HCVH
                                                           61
                                                                                                                                             121
                                                                                                                                                                                                                                                                                                                                RESULT 2
POLG_HCVH
                                                                                                                                                                                                                                                                                                                                                                                                q
                                                                                                                                                                                                                             ò
                                                         ò
                                                                                                   g
                                                                                                                                             ò
```

BORITO 305 305 N.LINKED (GLCNAC ) (PP BORITO 417 417 N.LINKED (GLCNAC ) (PP BORITO 418 418 418 N.LINKED (GLCNAC ) (PP BORITO 540 641 N.LINKED (GLCNAC ) (PP BORITO 540 N.LINKED 610 N.LINKED (GLCNAC ) (PP BORITO 540 N.LINKED 610 N.LINKED	<pre>uery Match 97.3%; Score 927; DB 1; Length 3011; sest Local Similarity 96.2%; Pred. No. 1.2e-80;</pre>
LY U39.  LY S29.  LY S29.  LY S29.  As collaboration way a service on its service on its service on its service on way a service on its service on way a service of the commercial service of the commer	) (POIENTIAL). Query) (POIENTIAL). Best
PROPRIETY PROCESSING AT HE PROPERSE BELONGS TO PETILIASE FAMILABITY: THE NA3 PROTEASE ALEMANCA AND ASSA LIST CONTENT AND ASSA NOT AND ASSA LIST CANADA AND ASSA LINCEPTOR IPRODODAS HACK NASA LINCEPTOR IPRODOSS HACK NASA LIN	234 234

```
Matches 164;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Query Match
Best Local
  This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                         Ö
                                                                                                GDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVCTRGVAKAVDFIPVESLETTM 180
                   09
                                                                                                                                                                                                                   01-APR-1993 (Rcl. 25, Created)
01-APR-1993 (Rcl. 25, Iast sequence update)
01-APR-1993 (Rcl. 25, Iast sequence update)
15-SEP-2003 (Rcl. 42, Iast annotation update)
16-Genome polyprotein (Contains: Capsid protein C (Core protein (P22);
Envelope glycoprotein El (GP32) (GP35); Envelope glycoprotein E2
(GP68) (GP70) (NS1); Protein P7; Nonstructural protein NS2 (P21)
(Rc 3 4 . 22. -); Proteassc/helicase NS3 (P70) (Hepacivirin)
(RC 3 4 . 21.98); Nonstructural protein NS4A (P4); Nonstructural protein NS4B (P27); Nonstructural protein NS5A (P65); Nonstructural protein NS5B (P66) (P70) (RNA-directed RNA polymerase) (EC 2.7.7.48)].
NINUSES, SSRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                   GTRT1ASPKGPVIQMYTNVDKDLVGWPAPQGSRSLTPCTCGSSDLYLVTRHADVIPVRRR
                  1 MAPITAYAQQTRGLLGCIITSLTGRDKNQVEGEVQIVSTAAQTFLATCINGVCWTVYHGA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (RNA)(N).
SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MRNA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY S29.
Indels
3;
4; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             EMBL; M84754; -; NOT_ANNOTATED_CDS.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     MEROPS; S29.001; ...
MEROPS; U39.001; ...
InterPro; IPR001410; DEAD.
Matches 175; Conservative
                                                                                                                                                                                                       STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    PDB; 1N64; 25-FEB-03
PDB; 1NS3; 08-APR-98
                                                                                                                                                                                                                                                                                                                                                   NCBI_TaxID=31645;
                                                                                                                                                       RS 1207
                                                                                                                                     181 RS 182
                                                                                                                                                                                                       POLG_HCVTW
                                                                                                                                                       1206
                                                        61
                                                                                               121
                                                                                                                                                                                                      a
                                                                                              ò
                                                                                                                                                  셤
                                   q
                                                        δ
                                                                          g
```

```
CELLULAR AMINOPEPTIDASE.
CORE PROTEIN (POTENTIAL).
MAJRIX PROTEIN (POTENTIAL).
MAJOR ENVELOPE PROTEIN E (POTENTIAL).
NONSTRUCTURAL PROTEIN NS1,FC2 (POTENTIAL).
PROTEASE/HELICASE NS3 (POTENTIAL).
PROTEASE/HELICASE NS3 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4A (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4B (POTENTIAL).
RNA-DIRECTED RNA POLYMERASE (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Polyprotein; Glycoprotein; Transferase; RNA-directed RNA polymerase; Core protein; Coat protein; Envelope protein; Helicase; ATP-binding; Transmembrane; Nonstructural protein; Hydrolase; Serine protease;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  REMOVED FROM CAPSID PROTEIN C BY THE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SIMILARITY).
SIMILARITY).
SIMILARITY).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  (POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       (POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Length 3010;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                327047 MW; AAD267D55CDFE215 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 RELAY SYSTEM (BY
RELAY SYSTEM (BY
RELAY SYSTEM (BY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Score 892; DB 1; L
Pred. No. 2.7e-77;
2; Mismatches 6;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           N-LINKED (GLCNAC,
N-LINKED (GLCNAC,
N-LINKED (GLCNAC,
N-LINKED (GLCNAC,
N-LINKED (GLCNAC,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  (GLCNAC.
(GLCNAC.
(GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ATP (POTENTIAL)
DECH BOX.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       N-LINKED (N-LINKED (N-LINK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            N-LINKED
N-LINKED
N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  N-LINKED (
N-LINKED (
N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 N-LINKED
N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       N-LINKED
                                                                                                                                                                | IPR00166; HCV_RGRP.
| IPR001065; RNA_pol_DS_PS.
| IPR007095; RNA_pol_DS_PS.
| IPR007095; RNA_pol_DS_PS.
| IPR007095; RNA_pol_DS_PS.
| IPR007094; RNA_pol_PSVIT.
| F01543; HCV_Core; I.
| F01542; HCV_Core; I.
| F01559; HCV_NS1; I.
| F01560; HCV_NS1; I.
| F01560; HCV_NS3; I.
| F01000; HCV_NS4; I.
| F01000; HCV_NS5; I.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CHARGE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CHARGE
                HCV_NS1.
HCV_NS3.
HCV_NS4a.
HCV_NS4a.
                                                                                                                                                    HCV_NS5a.
HCV_RdRP.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           93.68;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 2240
2529
2788
                   IPR002531; F IPR002518; H IPR004109; H IPR000745; H IPR001490; H
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Š.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            2041
2077
2240
2529
2788
3010
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SM00487;
                                                                                                                                                                                                                                                            Pfam; PF01543;
                                                                                                                                                                                                                                                                                                                                                                       Pfam; PF01538;
                                                                                                                                                                                                                                                                                                                                                                                                                       Pfam; PF01006;
                                                                                                                                                                                                                                                                                                                                                                                                                                             Pfam; PF01001;
Pfam; PF01506;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Pfam; PF00271;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Pfam; PF03998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      3D-structure.
INIT_MET
                                                                        InterPro;
InterPro;
InterPro;
                                                                                                                                                    InterPro;
InterPro;
                   InterPro;
InterPro;
                                                                                                                                                                                                            InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ACT_SITE
ACT_SITE
ACT_SITE
NP_BIND
SITE
CARBOHYD
                                                                                                                                                                                                                                   InterPro;
nterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ProDom; F
SMART; SA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        TRANSMEM
                                                                                                                                                                                                                                                                                                               Pfam;
                                                                                                                                                                                                                                                                                                                                        Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CHAIN
CHAIN
CHAIN
CHAIN
CHAIN
CHAIN
CHAIN
```

ö

Gaps

; 0

Indels

12; Mismatches

Conservative

Similarity

S

```
CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CHAIN
1026 LAPITAYAQQTRGLFGCIITSLfGRDKNQVEGEVÇVVSTATQSFLATCINGVCWTVYHGA 1085
                                                                                 GTRIIASPKGPVIQMYINVDKDLVGWPAPQGSRSLIPCTCGSSDLYLVTRHADVIPVRRR 120
                                                                                                                                                                121 GDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVCTRGVAKAVDFIPVESLETTM 180
1 MAPITAYAOQTRGLLGCTITSLTGRDKNOVEGEVQIVSTAAQTFLATGINGVCWTVYHGA 60
                                                                                                                                                                                                                                                                                                                                                                                                                                                        01-802 (Rel. 23, Last sequence update)
15-SEP-2003 (Rel. 42, Last annotation update)
16-SEP-2003 (Rel. 42, Last annotation (P22);
16-SEP-2003 (Rel. 42, Last annotation El (GP32) (GP35); Envelope glycoprotein E2
16-SEP-2003 (GP70) (NS1); Protein P7; Nonstructural protein NS4 (P27); Nonstructural protein NS4 (P47); Nonstructural protein NS5 (P55); Nonstructural protein NS5 (P55); Nonstructural protein NS5 (P65) (P70) (RM-4directed RNA polymerase) (EC 2.7.7.48);
16-SEP-2003 (SP30) (RM-2014) (R
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        the viral
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  X-RAY CRYSTALLOGRAPHY (2.4 ANGSTROMS) OF 1027-1215.
MEDLINE-97015088: PubMed-8861916:
LOVE R.A., Parge H.E., Wickcrsham J.A., Hostomsky Z., Habuka N.,
MOGMAN E.W., Adachi T., Hostomska Z.;
"The crystal structure of hepatitis C virus NS3 proteinase reveals a trypsin-like fold and a structural zinc binding site.";
Cell 87:331-342(1996).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Takamizawa A., Mori C., Fuke I., Manabe S., Murakami S., Fujita J., Onishi E., Andoh T., Yoshida I., Okayama H.;
"Structure and organization of the hepatitis C virus genome isolated
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS) OF 1027-1210 AND 1678-1691. MEDLINE-98227846; PubMed-9568891; Park Y., Li Y., Munshi S., Sardana V., Cole J.E., Sardana M., Steinkuehler C., Tomsi L., de Francesco R., Kuo L.C., Chen Z.; Complex of NS3 protease and NS4A peptide of BK strain hepatitis C
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Protein Sci. 7:837-847(1998).

-!- FUNCTION: THE SHALL PROTEINS NSZA, NSZB, NS4A AND NS4B ARE HYDROPHOBIC, SUGGESTING A POSSIBLE MEMBRANE-RELATED FUNCTION.

NS3 AND NS5 MAY PLAY A ROLE IN THE VIRAL RNA REPLICATION.

-!- CATALYTIC ACTIVITY: Hydrolysis of four peptide bonds in the vip precursor polyprotein, commonly with Asp or Glu in the P6 possition, Cys or Thr in Pl and Ser or Ala in Pl'.

-!- CATALYTIC ACTIVITY: N nucleoside triphosphate - N diphosphate.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 virus: a 2.2-A resolution structure in a hexagonal crystal form.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       MEDIINE-56235224, PubMed-8647104;
Borowski P., Heiland M., Oehlmann K., Becker B., Korneteky L.
"Non-structural protein 3 of hepatitis C virus inhibits
phosphorylation mediated by CAMP-dependent protein kinase.";
Bur. J. Biochem. 237:611-618(1996).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SEQUENCE FROM N.A.
MEDLINE-91140698; Pubmed-1847440;
                                                                                                                                                                                                                                                                                                                                                                                                                                           01-AUG-1992 (Rel. 23, Created)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Virol. 65:1105-1113(1991).
                                                                                                                                                                                                                                                                                                                                                                                                 STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SEQUENCE OF 1487-1500.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                from human carriers."
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               NCBI_TaxID-11105;
                                                                                                                                                                                                                                                                                         1206 RS 1207
                                                                                                                                                                                                                                                  181 RS 182
                                                                                                                                                                                                                                                                                                                                                                                               POLG_HCVBK
P26663;
                                                                                     19
                                                                                                                                                                                                                                                                                                                                                                           POLC_HCVBK
ID POLC_H
                                                                                                                                                                                                                                                                                                                                                                                                                       a
                                                                                                                      g
                                                                                                                                                                                                                                                    ö
                                   g
                                                                                   ò
                                                                                                                                                                  ò
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation. The European Bloinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CELLUIAR ANINOPEPIIDASE.
CAPSID PROTEIN C (POTENTIAL).
MATRIX PROTEIN (POTENTIAL).
MAJOR ENVELOPE PROTEIN E (POTENTIAL).
NONSTRUCTURAL PROTEIN NSI/E2 (POTENTIAL).
NONSTRUCTURAL PROTEIN NSI/E2 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS3 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4A (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4A (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Polyprotein; Glycoprotein; Transferase; RNA-directed RNA polymerase;
Core protein; Coat protein; Envelope protein; Helicase; ATP-binding;
Fransmembrane; Nonstructural protein; Hydrolase; Serine protease;
SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 RNA-DIRECTED RNA POLYMERASE (POTENTIAL)
                                                          PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MRNA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       REMOVED FROM CAPSID PROTEIN C BY THE
                                                                                                                                                   -:- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY $29.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CHARGE RELAY SYSTEM.
CHARGE RELAY SYSTEM.
CHARGE RELAY SYSTEM.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ATP (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            POTENTIAL.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Interpro; IPR001410; DEAD.
Interpro; IPR001241; HCV_core.
Interpro; IPR002521; HCV_core.
Interpro; IPR002521; HCV_core.
Interpro; IPR002519; HCV_env.
Interpro; IPR002518; HCV_MS1.
Interpro; IPR004109; HCV_MS3.
Interpro; IPR001409; HCV_MS3.
Interpro; IPR001409; HCV_MS4.
Interpro; IPR001409; HCV_MS4.
Interpro; IPR001409; HCV_MS4.
Interpro; IPR001409; HCV_MS5.
Interpro; IPR001090; HCV_MS5.
Interpro; IPR001090; HCV_MS5.
Interpro; IPR001099; RNA_pol_pS.PS.
Interpro; IPR001099; HCV_CORE; IPPfam; PF01539; HCV_CORE; IPPfam; PF01539; HCV_MS1; IPPfam; PF01001; HCV_MS3; IPPfam; PF01001; HCV_MS4b; IPPfam; PF01606; HCV_MS4b; IPPGAm; IPPGAm;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Pfam; PF00998; Viral_RdRP; 1.
ProDom; PD186062; HCV_NS1; 1.
SMART; SMO0487; DEXDC; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  EMBL; M58335; AAA72945.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          729
1006
1615
1862
2013
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           PDB: 1JXP; 14-JAN-98.
PDB: 1C2P; 15-NOV-00.
PDB: 1C2P; 15-NOV-00.
PDB: 1GX5; 08-NOV-99.
PDB: 1GX5; 09-APR-02.
PDB: 1GX6; 10-APR-02.
PDB: 1QUV: 26-JUN-00.
PDB: 00-MY; 20-APR-99.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1A10; 25-MAR-98
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     PIR; A38465; GNWVTC
PDB; 1A1Q; 25-MAR-9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ; S29.001;
: U39.001;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    116
192
384
730
1007
1616
1863
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                347
1083
1107
1165
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        3D-structure.
INIT_MET
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ACT_SITE
ACT_SITE
ACT_SITE
NP_BIND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    TRANSMEM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  MEROPS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 MEROPS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CHAIN
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                            01-40G-1992 (Rel. 23, Created)
01-40G-1992 (Rel. 23, Last sequence update)
01-40G-1992 (Rel. 23, Last sequence update)
01-60G-1992 (Rel. 41, Last annotation update)
Genome polyprotein [Contains capsid protein C (Core protein) (P22);
Envelope qlycoprotein El (GP32) (GP35); Envelope glycoprotein E2
(GP68) (GP70) (NSI); Protein P7: Nonstructural protein NS2 (P21)
(EC 3.4.22-); Protease/helicase NS3 (P70) (Hepacivirin)
(EC 3.4.21-98); Nonstructural protein NS5A (P4); Nonstructural protein
NS4B (P27); Nonstructural protein NS5A (P4); Nonstructural protein
NS5B (P66) (P70) (RM-4directed RNA Polymerase) (EC 2.7.7.48)].
Hepatitis C virus (isolate Japanese) (HCV)
Viruses; SSRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                             DISCUSSION OF SEQUENCE.

MEDLINE-91192160; PubMed-1849488;

KALO N., Hijikata M., Nakagawa M., Ootsuyama Y., Muraiso K.,

KALO N., Hijikata M., Nakagawa M., Ootsuyama Y., Muraiso K.,

Molecular structure of the Japanese hepatitis C viral genome.";

FEBS Lett. 280:325-328(1991).

I-PUNCTION: THE SHALL, PROTEINS NS2A, NS2B, NS4A AND NS4B ARE

HYDROPHOBIC, SUGGESTING A POSSIBLE MEMBRANE-RELATED FUNCTION.

NS3 AND NS5 MAY PLAY A ROLE IN THE VIRAL RNA REPLICATION.

I-CATALITIC ACTIVITY: Hydrolysis of four peptide bonds in the viral precursor polyprotein, commonly with Asp or Glu in the P6 position, Cys or thr in Pl and Ser or Ala in Pl'.

CATALITIC ACTIVITY: N nucleoside triphosphate + CATALITIC ACTIVITY: N nucleoside triphosphate - N diphosphate
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (RNA)(N).
SUBUNT: THE VIRION OF THIS VIRUS IS A NUCLECCAPSID COVERED BY LIPOPROTEIN ENVELOPE CONSISTS OF TWO PROTEINS:
PROTEIN M AND GLYCOPROTEIN E. THE NUCLECCAPSID IS A COMPLEX OF PROTEIN C AND MRNA.
SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY $29.
                                                                                                                                                                                                                                                       SEQUENCE FROM N.A.
MEDILTB-9108550; METON N. Hijikata M., Ooksuyama Y., Nakagawa M., Ohkoshi S.,
Kato N., Hijikata M., Ootsuyama Y., Nakagawa M., Ohkoshi S.,
Sugimura T., Shimotohno K.;
"Molecular cloning of the human hepatitis C virus genome from
"Molecular cloning of the human hepatitis C virus genome from
"Japanese patients with non-A, non-B hepatitis.",
Proc. Natl. Acad. Sci. U.S.A. 87:9524-9528(1990).
                                                          3010 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      HCV_capsid.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                HCV_core.
HCV_env.
HCV_NS1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   HCV_NS3.
HCV_NS3.
HCV_NS4a.
HCV_NS4b.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            EMBL; D90208; BAA14233.1; -. PIR; A39253; GNWVCJ. HSSP; P26663; JJXP.
                                                          STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      IPR001490;
IPR002868;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         InterPro; IPR001410;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         IPR002531;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         IPR000745;
                                                                                                                                                                                                                                 NCBI_TaxID-11116;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               MEROPS; $29,001;
MEROPS; U39.001;
1206 RS 1207
                                                        POLG_HCVJA
P26662;
                                                                                                                                                                                                                       Hepacivirus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    InterPro; ]
InterPro; ]
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            InterPro;
InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      InterPro;
InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             nterPro;
                                            POLG_HCVJA
                                  RESULT
                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      \dot{\circ}
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        121 GDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVCTRGVAKAVDFIPVESLETIM 19C
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1 MAPITAYAQQTRGLIGCIITSLIGRDKNOVEGEVOIVSTAAQTFLATCINGVCWTVYHGA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           GTRTIASPKGPVIQMYTNVDKDLVGWPAPQGSRSLTPCTCGSSDLYLVTRHADVIPVRRR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     0; Gaps
                                                                               (POTENTIAL).
                                                                                                                                                                                                            POTENTIAL)
                                               (POTENTIAL)
                                                          (POTENTIAL)
                                                                                                                 (POTENTIAL)
                                                                                                                               (POTENTIAL)
                                                                                                                                                    (POTENTIAL)
                                                                                                                                                                  (POTENTIAL)
                                                                                                                                                                                                                                    (POTENTIAL)
                                                                                                       (POTENTIAL)
                                                                                                                                          (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Length 3010;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ; Score 891; DB 1; Length 30; Pred. No. 3.4e-77; 15; Mismatches 5; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         327189 MW; F8422D5ECCFDF09C CRC64;
                                            (GLCNAC.
(GLCNAC.
                                                                               (GLCNAC.
(GLCNAC.
(GLCNAC.
(GLCNAC.
(GLCNAC.
(GLCNAC.
(GLCNAC.
                                                                                                                                                                                                  (GLCNAC.
            N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              93.58;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          89.08;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Local Similarity 89.0%
Les 162; Conservative
                                                                                                                                                                                                                                                                                                                                                                         1097
11120
11120
11120
11133
11134
11144
11161
11161
11161
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1186
1188
1197
1202
                                                                                                                                                                                                                                                         1047
1050
1063
1074
                                                                                                                                                                                                                                                                                                     1076
1081
1085
1087
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1688
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        3010 AA;
                                                                                                                                                                                                                                               1031
1039
10050
11058
11068
11082
11098
11108
11120
11120
11139
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        RS 182
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1146
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           61
                                                                                                                               CARBOHYD
CARBOHYD
CARBOHYD
                                                                                                                                                                 CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SEQUENCE
                                                                                                       CARBOHYD
                                                                                                                                                                                                                       CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Query Match
                                                                       CARBOHYD
CARBOHYD
                                               CARBOHYD
             CARBOHYD
                          CARBOHYD
                                     CARBOHYD
                                                             CARBOHYD
                                                                                              CARBOHYD
                                                                                                                                                                                                                                                                                                                                       TURN
STRAND
TURN
                                                                                                                                                                                                                                                                                                                                                                         STRAND
STRAND
STRAND
STRAND
STRAND
STRAND
STRAND
STRAND
STRAND
TURN
TURN
TURN
STRAND
HELLX
TURN
STRAND
HELLX
TURN
STRAND
                                                                                                                                                                                                                                                                    STRAND
                                                                                                                                                                                                                                                                                           STRAND
                                                                                                                                                                                                                                                                                                                   STRAND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              STRAND
                                                                                                                                                                                                                                               STRAND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            STRAND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Best Loca
Matches
                                                                                                                                                                                                                                                          HELIX
QQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
```

```
RESULT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            61 GTRTIASPKGPVIOMYTNVDKDLVGWPAPQGSRSLTPCTCGSSDLYLVTRHADVIPVRRR 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     GDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVCTRGVAKAVDFIPVESLETTM 180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1 MAPITAYAQQTRGLLGCIITSLTGRDKNQVEGEVQIVSTAAQTFLATCINGVCWTVYHGA 60
                                                                                                                                                                                                                                   CAPGID PROTEIN C (POTENTIAL).
MAJOR ENVELOPE PROTEIN (POTENTIAL).
MAJOR ENVELOPE PROTEIN E (POTENTIAL).
NONSTRUCTURAL PROTEIN NS1 (POTENTIAL).
NON-STRUCTURAL PROTEIN NS2 (POTENTIAL).
PROTEASE/HELICASE NS3 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4B (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4B (POTENTIAL).
RNA-DIRECTED RNA POLYMERASE (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           0; Gaps
                                                                                                                                                                                POTENTIAL.
CHARGE RELAY SYSTEM (BY SIMILARITY).
CHARGE RELAY SYSTEM (BY SIMILARITY).
CHARGE RELAY SYSTEM (BY SIMILARITY).
                                                                                                                                                                                                                                                                                                                                                                                                                                                      (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                    (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                 (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                              (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      DB 1; Length 3010;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ; Pred. No. 8.2e-77;
17; Mismatches 5; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 MW; AA993794F46DB185 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                               N-LINKED GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                             ATP (POTENTIAL).
DECH BOX.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       93.1%; Score 887;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          N-LINKED
N-LINKED
      rPro: IPR001650; Helicase_C.
rPro: IPR007095; RNA_pol_DS_PS.
rPro: IPR007094; RNA_pol_PSvir.
pro1543; HCV_capsid; 1.
pro1542; HCV_core; 1.
                                                                       Pfam: PF01560; HCV_NS1: 1.
Pfam: PF01580; HCV_NS1: 1.
Pfam: PF0207; HCV_NS3: 1.
Pfam: PF01001; HCV_NS4a: 1.
Pfam: PF01001; HCV_NS4b: 1.
Pfam: PF01506; HCV_NS5a: 1.
Pfam: PF001506; HCV_NS5a: 1.
Pfam: PF00271; hellcase_C: 1.
Pfam: PF00998; Viral_RGRP: 1.
ProDom: PD186062; HCV_NS1: 1.
IPRO02166; HCV_RdRP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                3010 AA; 327017
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 87.98;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           160; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          2240
2788
                                                                                                                                                                        SMART; SM00487; DEXDC;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Similarity
                   InterPro;
InterPro;
InterPro;
                                                                                                                                                                                                                                                                                                                                                                   ACT_SITE
NP_BIND
SITE
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CARBOHYD
CARBOHYD
CARBOHYD
            InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                             CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                             CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      121
                                                                                                                                                                                                                                                                                                                         CHAIN
TRANSMEM
ACT_SITE
ACT_SITE
                                                                                                                                                                                                                                                                                                                                                                                                                                   CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Local
                                                                                                                                                                                                                                                  CHAIN
                                                                                                                                                                                                                                                                                 CHAIN
                                                                                                                                                                                                                                                                                                    CHAIN
                                           εam;
                                                                                                                                                                                                                                       CHAIN
                                                                Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Matches
g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license afterement (See http://www.isb.sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                        MEDLINE-92295714; PubMed-1318627;
MEDLINE-92295714; PubMed-1318627;
MEDLINE-92295714; PubMed-1318627;
MEDLINE-92295714; PubMed-1318627;
MEDLINE-92295714; PubMed-1318627;
MEDLINE-92295714; PubMed-1318627;
Medecular cloning of hepatitis C virus genome from a single Japanese cariser: sequence variation within the same individual and among infected individuals.":

Virus Res. 23:39-53(1992);
Virus Res. 23:39-53(1992);
Virus Res. 23:39-53(1992);
C -: FUNCTION: THE SMALL PROTEINS NS2A, NS2B, NS4A AND NS4B ARE NS3 AND NS5 MAY PARD FOR THE VIRAL MEMBRANE-RELATED FUNCTION.
NS3 AND NS5 MAY PATA A ROLE IN THE VIRAL RNA REPLICATION.
NS3 AND NS5 MAY PATA A ROLE IN THE VIRAL RNA REPLICATION.
C -: CATANTIC ACTIVITY: Hydrolysis of four peptide bonds in the P6 position, Cys or Thr in P1 and Ser or Ala in P1.
C -: CATANTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
                                                                                                                                       01-APR-1993 (Rel. 25, Created)
01-APR-1993 (Rel. 25, Last sequence update)
01-APR-1993 (Rel. 25, Last sequence update)
01-Process (Rel. 42, Last as noncation update)
Genome polyprotein (Contains notation and protein (P22);
Envelope glycoprotein El (GP32) (GP35); Envelope glycoprotein E2
(GP68) (GP70) (NS1); Protein P7; Nonstructural protein NS2 (P21)
(EC 3.4.22.-); Protease/helicase NS3 (P70) (Hepacivirin)
(EC 3.4.21.98); Nonstructural protein NS48 (P4); Nonstructural protein
NS48 (P27); Nonstructural protein NS5A (P56); Nonstructural protein
NS5B (P66) (P70) (RNA-directed RNA pollymerase) (EC 2.7.7.48)].
Hepatitis C virus (isolate HC-JT) (HCV).
Viruses; SSRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              -1. SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS: PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MRNA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY S29.
                                                                                                        PRT; 3010 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      RNA_pol_DS_PS.
RNA_pol_PSvir.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                HCV_capsid.
HCV_core.
HCV_env.
HCV_NSI.
HCV_NS2.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     HCV_NS4a.
HCV_NS4b.
HCV_NS5a.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        EMBL; D11168; BAA01943.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Interpro; IPR002522; H
Interpro; IPR002521; H
Interpro; IPR002519; H
Interpro; IPR002511; H
Interpro; IPR002518; H
Interpro; IPR0004109; H
Interpro; IPR000445; H
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             PDB; 1A1Q; 25-MAR-98.
PDB; 1JXP; 14-JAN-98.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               InterPro; IPR001410;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             IPR007094;
                                                                                                                                                                                                                                                                                                                                                                                                        NCBI_TaxID-31642;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     MEROPS; S29.001;
MEROPS; U39.001;
1206 RS 1207
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            PIR; A45573;
                                                                                                     POLG_HCVJT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          InterPro;
                                                                                    POLG_HCVJT
```

II RS 1207

```
DEPT TO DEPT T
                            qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ô
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1026 LAPITAYAQQTRGLLGCIVTSLTGRDKNQVEGEVQVVSTATQSFLATCVNGVCWTVFHGA 1085
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   GDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVCTRGVAKAVDFIPVESLETIM 180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              09
                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ATTENDED FROM LANGUID FROM LINE CELULLAR AMINOBETIDASE.
CAPSID PROTEIN (POTENTIAL).
MATRIX PROTEIN (POTENTIAL).
MAJOR ENVELOPE PROTEIN E (POTENTIAL).
NONSTRUCTURAL PROTEIN NSI/E2 (POTENTIAL).
NON STRUCTURAL PROTEIN NSI/E2 (POTENTIAL).
NONSTRUCTURAL PROTEIN NSI/E2 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4A (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4B (POTENTIAL).
POTENTIAL).
POTENTIAL).
                 P Fram. PF01542; HCV_COTE; 1.

R Pfam: PF01542; HCV_COTE; 1.

R Pfam: PF01560; HCV_CNU; 1.

R Pfam: PF01560; HCV_NS2; 1.

R Pfam: PF01006; HCV_NS3; 1.

R Pfam: PF01006; HCV_NS4; 1.

R Pfam: PF01001; HCV_NS4; 1.

R Pfam: PF0027; HCV_NS4; 1.

R Pfam: PF0027; HCV_NS5; 1.

R Pfam:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      GTRIIASPKGPVIQMYINVDKDLVGWPAPQGSRSLIPCTCGSSDLYLVTRHADVIPVRRR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           MAPITAYAQQTRGLLGCIITSLTGRIXKNQVEGEVQIVSTAAQTFLATCINGVCWTVYHGA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       0; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SIMILARITY).
SIMILARITY).
SIMILARITY).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                    REMOVED FROM CAPSID PROTEIN C BY THE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (POTENTIAL).
(POTENTIAL).
(POTENTIAL).
(POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                (POTENTIAL)
(POTENTIAL)
(POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Length 3010;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    6: Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  INKED (GLCNAC. . .) (POT 94A1C77435D642BB CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CHARGE RELAY SYSTEM (BY
CHARGE RELAY SYSTEM (BY
CHARGE RELAY SYSTEM (BY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          92.8%; Score 884; DB 1;
88.5%; Pred. No. 1.6e-76;
ive 15; Mismatches 6.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             N-LINKED (GLCNAC,
N-LINKED (GL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ATP (POTENTIAL)
DECH BOX.
  HCV_capsid; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            326573
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Matches 161; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              645
2041
2077
2240
2529
2788
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1083
1107
1165
11237
1319
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      2529
2788
3010
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 RS 182
PF01543;
PF01542;
PF01539;
                                                                                                                                                                                                                                                                                                                                                                                                                                          3D-structure.
INIT_MET
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CARBOHYD
CARBOHYD
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           _
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ACT_SITE
ACT_SITE
ACT_SITE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      NP_BIND
SITE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    TRANSMEM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Local
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CHAIN
CHAIN
CHAIN
CHAIN
CHAIN
CHAIN
CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CHAIN
     ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ô
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    qq
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation he European Bioinformatics institute. There are no respirations on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to licensefisb-sib.ch).
                                                                                                                            01-AUG-1992 (Rel. 23, Created)
01-AUG-1992 (Rel. 23, Last sequence update)
01-AUG-1992 (Rel. 23, Last sequence update)
01-AUG-1992 (Rel. 41, Last annotation update)
Genome polyprotein [Contains: Capsid protein C (Core protein) (P22);
Envelope glycoprotein El (GP32) (GP35); Envelope glycoprotein E2
(GP68) (GP70) (NSI); Protein P7; Nonstructural protein NS2 (P21)
(EC 3.4.22-); Protease/hellcase NS3 (P70) (Hepacivirin)
(EC 3.4.21-98); Nonstructural protein NSA (P4); Nonstructural protein
NSAB (P57); Nonstructural protein NSA (P4); Nonstructural protein
NSAB (P66) (P70) (RNA-directed RNA pollymerase) (EC 2.7.7.48)].
Hepatitis C virus (isolate HC-J8) (HCV).
Viruses; SSRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           VITOLOGY 188:331-341(1992).

-!- FUNCTION: THE SMALL PROTEINS NS2A, NS2B, NS4A AND NS4B ARE HYDROPHOBIC, SUGGESTING A POSSIBLE MEMBRANE-RELATED FUNCTION.

NS3 AND NS5 MAY PLAY A ROLE IN THE VIRAL RNA REPLICATION.

-!- CATALYTIC ACTIVITY: Hydrolysis of four peptide bonds in the viral precursor polyprotein, commonly with Asp or Glu in the P6 postition, Cys or Thr in P1 and Ser or Ala in P1'.

-!- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SEQUENCE FROM N.A.
MEDIJUE-92230332; PubMed-1314459;
MEDIJUE-922303322; PubMed-1314459;
MEDIJUE-922303322; PubMed-1314459;
Pubmoto H., Kurai K., Okada S.-I., Yamamoto K., Lizuka H., Tanaka T.,
Fukuda S., Tsuda F., Mishiro S.;
"Full-length sequence of a hepatitis C virus genome having poor homology to reported isolates: comparative study of four distinct
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A LIPOPROTEIN BYNPELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
PROFIEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROFIEIN C AND MRNA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY S29.
                                                                     PRT; 3033 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  InterPro; IPR002868; HCV_NS5a.
InterPro; IPR002166; HCV_RdRP.
InterPro; IPR007095; RNL_DDL_DS_PS.
InterPro; IPR007094; RNA_pol_PSvir.
Pfam; PF01543; HCV_capsid; 1.
Pfam; PF01542; HCV_core: 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             HCV_capsid.
HCV_core.
HCV_env.
HCV_NSI.
HCV_NSI.
HCV_NSI.
HCV_NSI.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   HCV_NS4b
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               EMBL; D10988; BAA01761.1; -. PIR; A40250; GNWVJ8.
                                                                     STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                InterPro; IPR001410;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       IPR002519;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         IPR002518;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        PIR; A40250; GNWVJ8.
HSSP; P27958; 1HEI.
MEROPS; S29.001; -.
MEROPS; U39.001; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     THEI.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  NCBI_TaxID=11115;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          {RNA}(N).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Hepacivirus
                                                                  POLG_HCVJ8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             InterPro;
InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             genotypes.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                InterPro;
                                    POLG_HCV38
RESULT
```

```
POLG_HCVJ6
P26660;
                                                                                                                                                                                                                                                    Hepacivirus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  61 GTRTIASPKGPV1QMYINVDKDLVGWPAPQGSRSLTPCTCGSSDLYI,VTRHADVIPVRRR 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              121 GDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVCTRGVAKAVDFIPVESLETIM 180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1 MAPITAYAQQTRGLLGCIITSLTGRDKNQVEGEVQIVSTAAQTFLATCINGVCWTVYHGA 60
                                                                                                                                                                                                                        CAPSID PROTEIN C (FOTENIAL).
MATRIX PROTEIN (FOTENIAL).
MAJOR ENVELOPE PROTEIN E (POTENIAL).
NONSTRUCTURAL PROTEIN NSI (POTENIAL).
NONSTRUCTURAL PROTEIN NSI (POTENIAL).
NONSTRUCTURAL PROTEIN NSI (POTENIAL).
NONSTRUCTURAL PROTEIN NSA (FOTENIAL).
NONSTRUCTURAL PROTEIN NSAB (POTENIAL).
RNA-DIRECTED RNA POLYMERASE (POTENIIAL).
POTENIAL.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       0; Gaps
                                                                                                                                                      Polyprotein; Glycoprotein; Transferase; RNA-directed RNA polymerase; Core protein; Coat protein; Envelope protein; Holicase; ATP-binding; Transmembrane; Nonstructural protein; Hydrolase; Serine protease.
INIT_MET 1 CELLULAR AMINOPEPTIDASE.
                                                                                                                                                                                                                                                                                                                                                                                 CHARGE RELAY SYSTEM (BY SIMILARITY)
AIP (POTENTIAL)
N-LINKED (GLCNAC...) (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ; Score 714; DB 1; Length 3033;
; Pred. No. 3.4e-60;
26; Mismatches 27; Indels C
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1A173E7E3381FD1A CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            330177 MW;
Pfam; PF01539; HCV_env; 1. Pfam; PF01560; HCV_NS1: 1. Pfam; PF01580; HCV_NS2: 1. Pfam; PF02907; HCV_NS3: 1. Pfam; PF01006; HCV_NS4s; 1. Pfam; PF01001; HCV_NS4b; 1. Pfam; PF01506; HCV_NS5a; 1. Pfam; PF01506; HCV_NS5a; 1. Probom; PF0186062; HCV_NS1: 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         74.9%;
70.9%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Query Match
Best Local Similarity 70.99
Matches 129; Conscrvative
                                                                                                                                           SM00487; DEXDC;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Š
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             |:
|210 RT 1211
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            RS 182
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                         ACT_SITE
ACT_SITE
NP_BIND
                                                                                                                                                                                                                                                                                                                                                                                                                                                     SITE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                          TRANSMEM
ACT_SITE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CARBOHYL
                                                                                                                                           SMART;
                                                                                                                                                                                                                                                  CHAIN
CHAIN
CHAIN
CHAIN
CHAIN
                                                                                                                                                                                                                                    CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            a
    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the Buropean Bioinformatics institute. There are no restrictions on its use by non-profit institutions and so a sits content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1. SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY S29.
                                                                                                                                                                                                                                                                                                                                                                     A MEDILINE-9204440; PubMed-1658196;
A Okamoto H., Okada S.-I., Sugiyama Y., Kurai K., Lizuka H.,
A Machida A., Miyakawa Y., Mayumi M.;
Machida A., Miyakawa Y., Mayumi M.;
Machida A., Miyakawa Y., Mayumi M.;
Mucleotide sequence of the genomic RNA of hepatitis C virus isolated
from a human carrier: comparison with reported isolates for conserved
T don a human carrier: comparison with reported isolates for conserved
T don. Virol. 72:2697-2704 (1991).
J. Gen. Virol. 72:2697-2704 AND NS4A AND NS4B AND
C -: FUNCTION: THE SMALL PROTEINS NS2A, NS2B, NS4A AND NS4B ARE
HYDROPHOSIC. SUGGESTING A POSSIBLE MEMBRANE-RELATED FUNCTION.
NS3 AND NS5 MAY PLAY A ROLLE IN THE VIRAL RNA REPLICATION.
NS3 AND NS5 MAY PLAY A ROLLE IN THE VIRAL RNA REPLICATION.
DECURSOR DOLYPOCHEDIN, commonly with Asp or Glu in the P6
position, Cys or Thr in Pl and Ser or Ala in Pl'.
C --- CATALYTIC ACTIVITY: N nucleoside triphosphate - N diphosphate +
                               01-AUG-1992 (Rel. 23, Created)
01-AUG-1992 (Rel. 23, Last sequence update)
01-AUG-1992 (Rel. 23, Last sequence update)
01-AUG-1992 (Rel. 24, Last annotation update)
Genome polyprotein (Contains: Capsid protein C (Core protein) (P22);
Envelope glycoprotein El (GP32) (GP35); Envelope glycoprotein E2
(GP68) (GP70) (NS1); Protein P7: Nonstructural protein NS2 (P21)
(EC 3 4.22.-); Protease/helicase NS3 (P70) (Hepacivirin)
(EC 3 4.21.98); Nonstructural protein NS5A (P66); Nonstructural protein NS5A (P66) (P70) (RNA-directed RNA polymerase) (EC 2.7.7.48)].
NS5B (P66) (P70) (RNA-directed RNA polymerase) (EC 2.7.7.48)].
Hepatitis C virus (isolate HC-J6) (HCV).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       (RNA)(N).
SUBGNIT: THE VIRION OF THIS VIRUS IS A NUCLECCAPSID COVERED BY A
LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
PROTEIN M AND GLYCOPROTEIN E. THE NUCLECCAPSID IS A COMPLEX OF
PROTEIN C AND MRNA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         InterPro; IPR002519; HCV_env
InterPro; IPR002519; HCV_env
InterPro; IPR002519; HCV_NS;
InterPro; IPR004109; HCV_NS3.
InterPro; IPR004109; HCV_NS4.
InterPro; IPR001490; HCV_NS4.
InterPro; IPR001490; HCV_NS5a.
InterPro; IPR002166; HCV_NSSa.
InterPro; IPR002166; HCV_NSP.
InterPro; IPR001099; Hellcase_C.
InterPro; IPR001099; NA_POJ_DS_PS.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        HCV_capsid
HCV_core.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Pfam; PF01543; HCV_capsid; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         EMBL; D00944; BAA00792.1; -.
STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             PIR; JQ1303; JQ1303.
HSSP; P27958; 1HEI.
MEROPS; S29.001; ".
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  InterPro; IPR001410;
InterPro; IPR002522;
InterPro; IPR002521;
                                                                                                                                                                                                                                                                                                                                            [1]
SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                  NCBI_TaxID-11113;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  PF01560;
PF01538;
```

```
PA4019
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   HHOA_ARATH
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Best Loc
Matches
      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     GTRT1ASPKGPVIQMYTNVDKDI,VGWPAPQGSRSLTPCTCGSSDI,YLVTRHADVIPVRRR 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      GDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVCTRGVAKAVDFIPVESLEITM 180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     09
                                                                                                               1 MAPITAYAQQTRGLLGCIIISLTGRDKNQVEGEVQIVSTAAQTFLATCINGVCWTVYHGA
                                                                                                                                                                                                                                                                                                                                   CHARGE RELAY SYSTEM (BY SIMILARITY).
CHARGE RELAY SYSTEM (BY SIMILARITY).
CHARGE RELAY SYSTEM (BY SIMILARITY).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            (POTENTIAL). (POTENTIAL). (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (POTENTIAL).
(POTENTIAL).
(POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ..
                                                                                                                                                                                                                                                                                                                                                                                                                                                         (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Length 3033;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         26; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              MW; F957F5C1A273BE9E CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                          N-LINKED GLCNAC
N-LINKED GLCNA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Score 712; DB 1;
Pred. No. 5.3e-60;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                 ATP (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       29; Mismatches
                                                                                                                                                                                                                                                                                                                          POTENTIAL.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                 DECH BOX
Pfam; PF02907; HCV_NS3; 1.
Pfam; PF01006; HCV_NS4s; 1.
Pfam; PF01001; HCV_NS4b; 1.
Pfam; PF01506; HCV_NS5a; 1.
Pfam; PF00271; helicase_C; 1.
Pfam; PF00998; viral_RGRP; 1.
ProDom; PD186062; HCV_NS1; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          74.78:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              329165
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  649
1091
2038
2811
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ¥
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         al Similarity
127; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1210 RS 1211
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               RS 182
                                                                                                                                                                                                                                                                                                                                                                                                                                           CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      121
                                                                                                                                                                                                                                                                                                                                                                   ACT_SITE
NP_BIND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Query Match
Best Local
                                                                                                                                                                                                                                                                                                                                      ACT_SITE
                                                                                                                                                                                                                                                                                                                                                                                                             CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                              CARBOHYD
                                                                                                                                                                                                                                                                                                                          FRANSMEN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
    οŽ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
```

209 AM

PRT;

STANDARD;

PAAD\_PSEAE ID PAAD\_PSEAE AC Q9HX08;

RESULT 9

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss institute of Bioinformatics and the EMBL outstation the European Bioinformatics institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       -----YLVTRHADVIPVRRRGDSRGSLLSPR--PIS-----YLKGSSGGPLLCPA 148
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           8 AQQTRGLLGCIITSLTGRDKNQVEGEVQ-IVSTAAQTFLATCINGVCWTVYHGAGTRTIA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  17 AQYGLRLLDCLV------GEEREVHFLISKAAQLVMAT------ETDVA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SPKGP-----CICGSSDL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          HHOA_ARATH STANDARD; PRT; 321 AA.
09SEL7; 049507;
16-0CT-2001 (Rel. 40, Created)
15-SEP-2003 (Rel. 40, Last sequence update)
15-SEP-2003 (Rel. 42, Last annotation update)
Protease HhoA, chloroplast precursor (EC 3.4.21.-).
HHOA OR AT4G18370 OR F28312.31
Arabidopsis thaliana (Mouse-ear cress).
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
                                                                                                                                                                                                                                                                                                                                               STRAIN-ATCC 15692 / PAO1;
STRAIN-ATCC 15692 / PAO1;
MEDLINE-2043337 PubMed-10984043;
MEDLINE-2043337 PubMed-10984043;
Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warrener P., Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M. Garber R.L., Goltry L., Tolentino E., Westbrock-Wadman S., Yuan Y., Brody L.L., Coulter S.N., Folger K.N., Kas A., Larbig K., Lim R.M., Brody L.L., Spencer D.H., Wong G. K.-S., Wu Z., Paulsen I.T., Reizer J., Saler M.H., Hancock R.E.W., Lory S., Olson M.V.;
'Complete genome sequence of Pseudomonas aeruginosa PAO1, an
                                                                                                                                                                                                                      Bacteria; Proteobactería; Gammaproteobacteria; Pseudomonadales;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  InterPro; IPR003382; Flavoprotein.
Pfan; PF0341; Flavoprotein; 1.
Hypotherical protein; Lywes: Decarboxylase; Complete proteome.
SEQUENCE 209 AA; 22367 MW; 01FD081CC495D3F6 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 opportunistic pathogen.":
Nature 406:959-964(7000).
-i. SIMILARITY: BELONGS TO THE POLYPRENYL P-HYDROXYBENZOATE PHENYLACRYLIC ACID DECARBOXYLASES FAMILY.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        / Match 9.1%; Score 87; DB 1; Length 209; Local Similarity 26.2%; Pred. No. 0.28; hes 55; Conservative 18; Mismatches 61; Indels
16-OCT-2001 (Rel. 40, Created)
16-OCT-2001 (Rel. 40, Last sequence update)
16-OCT-2001 (Rel. 40, Last annotation update)
Probable aromatic acid decarboxylase (EC 4.1.1.-).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             149 GHAVGIFRAAVCTRGVAKAVDFIPVESLET 178
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          --APGFYHQ---PQSVEDLVDFVVARILNT 189
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  EMBL; AE004818; AAG07406.1; -.
                                                                                                                                                                                                                                     Pseudomonadaceae; Pseudomonas NCBI_TaxID-287;
                                                                                                                                                                                    Pseudomonas aeruginosa
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       H83144; H83144.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SSS SE H H H H H H H
```

```
34691 MW;
                                                                                                                                     EMBL; AF114386; AAF24060.1; -
                                                                                                                                                                                                                                                                                            Pfam; PF00089; trypsin; 1.
PRINTS; PR00834; PROTEASES2C.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  8.8%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        48; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Homo sapiens (Human)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               321 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                NCBI_TaxID-9606;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           AAMP_HUMAN
Q13685;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ACT_SITE
ACT_SITE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Query Match
                                                                                                                                                                                                                                                                                                                                                                                                      TRANSIT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Local
                                                                                                                                                                                                                                                                                                                                                                                                                                                              DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 AAMP_HUMAN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Best
  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Dp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ô
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its
                                                                                                                                                                                                                                                                                                          RAY MEDLINE-2008348 Stickems w. Estian K.-D. Terryn N.,
RA MEDLINE-2008348 Stickems w. Estian K.-D. Terryn N.,
RA POhl T. Ducsterboeft A., Stickems w. Estian K.-D. Terryn N.,
RA POHL T. Ducsterboeft A., Stickems w. Estian K.-D. Terryn N.,
RA Langlariner M. de Simone V. Obermaier B. Mache R., Mueller M.
RA Weicheslaartner M. de Simone V. Obermaier B. Mache R., Mueller M.
Reichert B., Portetail D., Perez-Alonso M., Boutry M., Bancroft I.,
Racken M., Welligdomenech P. Watson M., Schmidthein T.
Racken M., Welligh B., Bilham L., Robben J.,
Racken M., Welligh B., Bilham L., Robben J.,
Racken M., Welligh B., Bilham L., Robben J., Andenbussche F.,
Racken M., Welligh B., Bilham L., Robben J., Andenbussche F.,
Racken M., Welligh B., Bilham L., Robben J., Andenbussche F.,
Racken M., Welligh B., Bilham L., Robben J., Noretter P.,
Racken M., Welligh B., Welligh M., Lennard W., Mcetter P.,
Racker A., Buysabaett C., Glelen J., Villarroel R., De Ceronta A.,
Robitson B., Robers J., Cronin A., Ouall W., McLay K., Mayes R.,
RA Van Mondau W., Robers J., Cronin A., Ullarroel R., Derbnert T.-H.,
RA Dose S., de Haan M., Marster M., Crimm M., Lochmert T.-H.,
Rober C., Fuchs W., Fartmann B., Granderath K., Dauner D., Herzl A.,
Rober C., Fuchs W., Edman M., Bargue M., Parons D., Jesse T.,
RA Massenct O., Quigley F., Clabaul G., Muendlen A., Febber R.,
Schnabl S., Hillor R., Schnife W., Lecharry A., Cascuberta E.,
Ribbons T., Weber N., Vandebobl M., Bargues M., Pitch J.,
Racken L., Schwarz S., Scholler C., Wonfort A., Catcol F., Spieck I.,
Racken L., Schwarz S., Scholler P., Heber S., Francs P., Belake C.,
Racken L., Schwarz S., Scholler P., Heber S., Francs P., Belake C.,
Racken L., Schwarz S., Scholler P., Heber S., Francs P., Barnell L., Bedrid M., Strong C., Mullick S., Scholler M., Martiner S., Scholler M., Martiner S., Scholler M., Martiner S., Scholler M., Basegava A., Hanced R., Hiller L.,
Ramer J., Follon L., Mardis E., Marchen M., Wither P., Bantley D., Prelich S., McChan M., Johnson D.,
M. Ali J., Bergdoff A., Jones K., Drone 
                                                                                                         Lensch M.H.A., Sokolenko A., Herrmann R.G.;
"Identification and characterization of the chloroplast HhoA protease."
a homolog to the bacterial periplasmic protease HboA.";
submitted (DEC-1998) to the EMBL/GenHank/DDBJ databases.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CAUTION: Ref.2 sequences differ from that shown due to erroneous gene model prediction. AT4G18370 and AT4G18375 were originally fused into a single gene.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Submitted (JUL-2001) to the SWISS-PROT data bank.
-1- SUBCELLULAR LOCATION: Chloroplast; within the thylakoid lumen.
-1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY $2C.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SEQUENCE OF 72-82; 96-110; 150-159; 178-211 AND 306-320.
Schubert M., Peterson U., Funk C., Haas B., Schroeder W.P.,
Kieselbach T.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   chloroplast lumen from Arabidopsis thaliana.";
eurosids II; Brassicales; Brassicaceae; Arabidopsis
                                                                                                                                                                                                                                                                     STRAIN-cv. Columbia;
MEDLINE-20083488; PubMed-10617198;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Nature 402:769-777(1999).
                                                              [1]
SEQUENCE FROM N.A.
                                                                                                                                                                                                                        [2]
SEQUENCE FROM N.A.
                      NCBI_TaxID=3702;
```

```
use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Ξ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            114 VIPVRRRGDSRG------SLLSPRPISYLK------GSSGGPLLCPA 148
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            218 AI-----GNPYGYENTLTIGVVSGLGREIPSPNGKSISEAIQTDADINSGNSGGPLLDSY 272
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          117 LTDEENGKIEGTG-----SGFVWDKLGHIVTNYHVIAKLATDQFGLQRCK 161
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       -i-TISSUE SPECIFICITY: EXPRESSED IN BLOOD VESSELS. STRONGLY EXPRESSED IN BLOOD VESSELS. STRONGLY EXPRESSED OLD N ENDOPHELIAL CELLS, CYTOTROPHOBLASTS, AND POORLY DIFFERENTIATED COLON ADENOCARCINOMA CELLS FOUND IN LYMPHATICS.
-i-SIMILARITY: Contains W WD repeats.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                MEDINE-92562124; PubMed-7743515;
MEDINE-92562124; PubMed-7743515;
MEDINE-92562124; PubMed-7743515;
Gallardo J.A., Liotta L.A.;
Gallardo J.A., Liotta L.A.;
"Identification of a new immunoglobulin superfamily protein expressed in blood vessels with a heparin-binding consensus sequence."; Cancer Res. 55:2140-2149(1995).
-!- FUNCTION: MAY HAVE A FUNCTION IN MIGRATING CELLS.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       59 -----GAGTRTIASPKGPVIQMYINVDKDLVGWPAPQGSRSLTPCTCGSSDLYLVTRHAD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Gaps
                                                                                                                                                                                                                                                                                                                           Hydrolase; Serine protease; Chloroplast; Thylakoid; Transit peptide.
TRANSIT 1 26 CHLOROFLAST (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Eukaryota, Metazoa, Chordata, Craniata, Vertebrata, Buteleostomi,
Mammalia: Eutheria: Primates, Catarrhini, Hominidae, Homo.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           82;
                                                                                                                                                                                                                                                                                                                                                                                                                                                          CHARGE RELAY SYSTEM (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CHARGE RELAY SYSTEM (POTENTIAL)
CHARGE RELAY SYSTEM (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Score 84; DB 1; Length 321; Pred. No. 0.88;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           60; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           22 LIGRDKNQVEGEVQIVSTAAQTFLATCINGVCW-----TVYH-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      68DB81E0BD27A7A7 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              149 GHAVGIFRAAVCTR---GVAKAVDF-IPVESLETIM 180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           273 GHTIGV-NTATFTRKGSGMSSGVNFAIPIDIVVRTV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    28-FEB-2003 (Rel. 41, Last annotation update)
Angio-associated migratory cell protein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       15-JUL-1998 (Rel. 36, Created)
15-JUL-1998 (Rel. 36, Last sequence update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 G (IN REF
                                                                                                                                                                                                                                                                                                                                                                                                       PROTEASE HHOA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        26; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SEQUENCE FROM N.A., AND PARTIAL SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                  THYLAKOID.
                                                                                                                                               EMBL; AL021710; CAA16717.1; ALT_SEO. EMBL; AL161548; CAB78839.1; ALT_SEO. MEROPS; SO1.279; -- Interpro; IPR001940; Protease2C. Interpro; IPR001254; Ser_protease_Try. Ffam; PF00089; trybsin; 1
                                                                                                                                                                                                                                                                                                                                                                                                                               POLY -GLU.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          PRT;
```

```
spirochete.";
                                                                                                                                                                                                                                                                                                                        Query Match
Best Local Simi
Matches 44; (
                                                                                                                                                                                                                                                                                        DOMAIN
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              257
                                                                                                                                                                                                                                           DOMAIN
                                                                                                                                                                                                                                                      NI AMOC
                                                                                                                                                                                                            SIGNAL
                                                                                                                                                                                                                                                               DOMAIN
                                                                                                                                                                                                                    CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            DEG1_ARATH
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 RESULT 13
    Q
                                                                                                                                                                                                                                                                                                                                                                                                                οy
                                                                                                                                                                                                                                                                                                                                                                                                                                      a
                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the Buropean Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                           6
                                                                                                                                                                                                                                                                                                                                                                                                                                      240
                                                                                                                                                                                                                                                                                                                                                                                                                54 WIVYHGAGIRIIASPKGPVIQMYINVDKDLVGWPAPQGSRSL----TPCTCGSSDLYLV 108
                                                                                                                                                                                                                                                                                                                                                                                                                                                           109 TRHADVIPVRRR---GDSRGS----LLSPRPISYLKGSSG--GPLLCPA------ 148
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                241 ----VLPDGKRAVVGYEDGTIRIWDLKQGSPIHVLKGTEGHQGPLTCVAANQDGSLILT 295
                                                                                                                                                                                                                                                                                                                                                                                                                                     197 WMEWH-----PRAPVLLAGT-ADGNTWMWKVPNGDCKTFQGPNCPATCGR-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Fraser C.M., Norris S.J., Weinstock G.M., White O., Sutton G.G., Dodgon R., Gwinn M., Hickey E.K., Clayton R., Ketchum K.A., Sodergren E., Hardham J.M., McLeod M.P., Salzberg S., Peterson J., Khalak H., Richardson D., Howell J.K., Chidambaram M., Utterback T., McDonald L., Artiach P., Bowman C., Cotton M.D., Fujii C., Garland S. Hatch B., Horst K., Roberts K., Sandusky M., Weidman J., Smith H.O., Venter J.C.;
                                                                                                                                                                                                                                                                                                                                                                                            Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Bacteria: Spirochaetes: Spirochaetales: Spirochaetaceae: Treponema
NCBI_TaxID=160;
                                                                                                                                                                                                                                                                                                                                                                                           64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 "Complete genome sequence of Treponema pallidum, the syphilis
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ------GHAVGIFR----AAVCTRGVAKAVDFIPVESL 176
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     296 GSVDCQAKLVSATTGKVVGVFRPETVASQPSLGEGESESNSVESL 341
                                                                                                                                                                                                                                                                                                                                                                     DB 1; Length 452;
                                                                                                                                                                                                                                                                                                                                                                                           Indels
                                                                                                                                                                                                                                HEPARIN-BINDING (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                               DA1413D25EB236C0 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                           47;
                                                                                                                         MIM; 603488; ... Fiheparin binding activity; TAS. GO:0002201; Fiheparin binding activity; TAS. Thaterpro; IPR01680; WD40.
Fina: PF00400; WD40; 8.
SWART: SM00320; WD40; 8.
PROSITE: PS00678; WD_REPEATS_1; 1.
PROSITE: PS50082; WD_REPEATS_2; 6.
PROSITE; PS50294; WD_REPEATS_REGION; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         16-OCT-2001 (Rel. 40, Created)
16-OCT-2001 (Rel. 40, Last sequence update)
16-OCT-2001 (Rel. 40, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                        Conservative 13; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     485 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Hypothetical lipoprotein TP0136 precursor TP0136.
                                                                                                                                                                                                                                                                                                                                                                    8.6%; Score 82; 25.3%; Pred. No.
                                                                                                                                                                                                                                          POLY-GLU.
WD 1.
WD 2.
WD 3.
WD 4.
WD 5.
WD 6.
WD 6.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            STRAIN-N1chols;
MEDLINE-98332770; PubMed-9665876;
                                                                                                                                                                                                                                                                                                                                               49015 MW:
                                                                                         EMBL; M95627; AAA6889.1; -. PIR; I39383; I39383.
Genew; HGNC:18; AAMP.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      STANDARD;
                                                                                                                                                                                                                               18
77
138
180
220
261
306
363
                                                                                                                                                                                                                     Repeat; WD repeat.
                                                                                                                                                                                                                                                                                                                                              452 AA;
                                                                                                                                                                                                                                                                                                                                                                    Ouery Match
Best Local Similarity
Matches 42; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Y136_TREPA
083172;
                                                                                                                                                                                                                                                                                                                                               SECUENCE
                                                                                                                                                                                                                                          DOMAIN
REPEAT
                                                                                                                                                                                                                                                               REPEAT
REPEAT
                                                                                                                                                                                                                                DOMAIN
                                                                                                                                                                                                                                                                                      REPEAT
                                                                                                                                                                                                                                                                                                  REPEAT
                                                                                                                                                                                                                                                                                                                        REPEAT
                                                                                                                                                                                                                                                                                                             REPEAT
                                                                                                                                                                                                                                                                                                                                    REPEAT
 q
                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            q
                                                                                                                                                                                                                                                                                                                                                                                                                  õ
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMEL cutstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entitles requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               78 NVDKDLVG-----WPAPQGSRSLIPCTC-----GSSDLYLVTRHADVIP-----VR 118
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         141 T----LVGGTSKPFWLVPGGTGNNGNCGCGGGGGSSSSSSCIHIWLVPGGTGNNGNCG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          23 TGRDKNQVEGEVQIVSTAAQTFLATCI--NGVCWTVYHGAG---TRTIASPKGPVIQMYT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
eurosids II; Brassicales; Brassicaceae; Arabidopsis.
Science 281:375-388(1998).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SEQUENCE FROM N.A., AND CHARACTERIZATION.
MEDIINE-98175982; PubMed-9507020;
Itzhaki H., Naveh L., Lindahl M., Cook M., Adam 2.;
Identification and characterization of DegP, a serine protease associated with the luminal side of the thylakoid membrane.";
J. Biol. Chem. 273:7094-7098(1998).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             HYPOTHETICAL LIPOPROTEIN TP0136. N-ACYL DIGLYCERIDE (POTENTIAL).
                                                                   (Potential).
-!- SIMILARITY: BELONGS TO THE TP013X FAMILY OF LIPOPROTEINS.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       8.6%; Score 82; DB 1; Length 485; 24.2%; Pred. No. 2.2; tive 13; Mismatches 65; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 DEGLARATH STANDARD; PRT; 437 AA.
022609; OSLK85;
16-OCT-2001 (Rel. 40, Last sequence update)
16-OCT-2001 (Rel. 40, Last sequence update)
28-FEB-2003 (Rel. 41, Last annotation update)
Protease Do-like 1, chloroplast precursor (EC 3.4.21.-).
BEGPI OR DEGP OR AT3627925 OR KIGNIZ.18.
Arabidopsis thallana (Mouse-ear cress).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                C7A4CEEDC7DC5CED CRC64;
                                                                                                                                                                                                                                                                                                                                                                          EMBL; AE001199; AAC65137.1; ALT_INIT.
TIGR; TP0136; --
Hypothetical protein; Lipoprotein; Membrane; Signal;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       GLY/SER-RICH.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  GLY/SER-RICH.
GLY/SER-RICH.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         POLY-SER.
POLY-SER.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                48984 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 23
485
24
178
210
267
327
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SEQUENCE FROM N.A. STRAIN-CV. Columbia;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    444
485 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Complete proteome.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        NCBI_TaxID-3702;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       152 VG 153
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           GG 258
```

```
210 PK--NKIRPIPVGVSADILVGQKVFAIGNPFGLDHTLTTGVISGLRREIS--SAATGRPI 265
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Pancreas; Digestion; 3D-structure.
ACTIVATION PEPTIDE.
PROPROTEINASE E.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             s_1
                                                                                                                                            (Rel. 37, Last sequence update)
(Rel. 42, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  -!- SUBCELLULAR LOCATION: EXTRACE!lular.
-!- TISSUE SPECIFICITY: Pancreas.
-!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY (PDE): 1FON; 14-OCT-96.
PDE): 1PT; 27-JAN-97.
MEROPS; S01.983; -.
                                                                                                            253 A.A
                                                                                                                                                                                                                                                                                                                                            [2]
SEQUENCE OF 14-253, AND DISULFIDE BONDS.
MEDLINE-86220198; PubMed-3519215;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       InterPro; IPR001314; Chymotrypsin.
InterPro; IPR001254; Ser_protease_Try.
                                                                                                                                                                                                                                                                                                                                                                                                                                                  X-RAY CRYSTALLOGRAPHY (1.7 ANGSTROMS).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Pfam; PF00089; trypsin; 1.
PRINTS; PR00722; CHYMOTRYPSIN.
SMART; SM00020; Tryp_SPC: 1.
PROSITE; PS50240; TRYPSIN_DOM; 1.
PROSITE; PS00134; TRYPSIN_HIS; 1.
PROSITE; PS00135; TRYPSIN_EER; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                             Eur. J. Biochem. 157:91-99(1986).
                                                                                                                                  (Rel. 09, Created)
                                                                                                                                                                                                                                                                                                                                                                                                                   procarboxypeptidase A complex
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Serine protease homolog;
                                                                                                           STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               253
57
103
206
187
227
                                                                                                                                                                                                                                       NCBL_TaxID-9913;
                                                                                                                                                                                                                                                                                                Pulgserver A.;
                                                                                                                                01-NOV-1988 (15-DEC-1998 (15-SEP-2003 (
                                                                                                           CAC3_BOVIN
P05805;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        DISULFID
DISULFID
DISULFID
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           DISULFID
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  PROPEP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CHAIN
                                                                                                           ö
                                                                                                                                                                                                                                                This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its was by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      7
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          89 PQGSRSLTPCTCGSSDLYLV------TRHADVIPVRRRGDSRGSLLSPRPI 133
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            --- TNVDKDLVGW-----PA 88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CHARGE RELAY SYSTEM (POTENTIAL).
CHARGE RELAY SYSTEM (POTENTIAL).
CHARGE RELAY SYSTEM (POTENTIAL).
HSPPSGLSNST -> SSTLFLHSPPSSHL (IN REF. 2).
V -> I (IN REF. 2).
P -> S (IN REF. 2).
G -> R (IN REF. 2).
LL -> HF (IN REF. 2).
LL -> WF (IN REF. 2).
LL -> WF (IN REF. 2).
L -> V (IN REF. 2).
L -> V (IN REF. 2).
C -> D (IN REF. 2).
C -> D (IN REF. 2).
C -> C (IN REF. 2).
                                                                                                                                                                 ENZYME REGULATION: INHIBITED BY PHENYLMETHYLSULFONYL FLUORIDE AND
                                                                                           Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ransit peptide; Chloroplast; Thylakoid. CHLOROPLASI (POTENTIAL).
            Kaneko T., Katoh T., Sato S., Nakamura A., Asamizu E., Tabata S.; "Structural analysis of Arabidopsis thaliana chromosome 3. 11. Sequence features of the 4.251,695 bp regions covered by 90 Pl. TAC
                                                                                                                                                                                      -1 - SUBCELLULAR LOCATION: BOUND TO LUMINAL SIDE OF THE THYLAKOID
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     55;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           8.4%; Score 80.5; DB 1; Length 437;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         __
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     55; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1497BIAB3F5FF2A4 CRC64;
                                                                                                                                                                                                           -1- INDUCTION: By heat shock.
-1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S2C.
-1- SIMILARITY: Contains 1 PDZ/DHR domain.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 PROTEASE DO-LIKE 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            SERINE PROTEASE.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Pred. No.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     THYLAKOID
                                                                                                                                                                                                                                                                                                                                                                                                           Interpro; IPR001940; Protease2C,
Interpro; IPR001254; Ser_protease_Iry.
                                                                                                                                                                                                                                                                                                                                                          EMBL; AP000371; BAB02539.1; -
EMBL; AP001302; BAR02539.1; JOINED.
MEROPS; S01.279; -
MEDLINE-20363099; PubMed-10907853
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  18;
                                                                                                                                                                                                                                                                                                                                                    EMBL; AF028842; AAC39436.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           56 VYHGAGTRTIASPKGPVIQMY-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      46213 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                          PRINTS; PR00834; PROTEASES2C.
SMART; SM00228; PDZ; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                PS50106; PDZ; 1,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      25.68;
                                                                                                                                                                                                                                                                                                                                                                                                                                  Pfam; PF00595; PDZ; 1.
Pfam; PF00089; trypsin; 1
                                                                                                                                                                                                                                                                                                                                                                                                  InterPro; IPR001478; PDZ.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Conservative
                                                           DNA Res. 7:217-221(2000).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   103
437
321
421
171
201
280
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        36
54
60
64
69
381
                                                                                                                                                                            O-PHENANTHROLINE.
                                                                      [3]
SEQUENCE OF 104-118.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Best Local Similarity
Matches 44; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      437 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 104
                                              and BAC clones.
                                                                                                                                                                                                  MEMBRANE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Hydrolase;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   DOMAIN
ACT_SITE
ACT_SITE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ACT_SITE
CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  PROSITE;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   TRANSIT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         TRANSIT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
```

```
"Autolysis of proproteinase E in bovine procarboxypeptidase A ternary complex gives rise to subunit III."; FEBS Lett. 277:37-41(1990).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 MEDLINE-94222022; PubMed-8168476; Pignol D., Gaboriaud C., Michon T., Kerfelec B., Chapus C., Fontecilla-Camps J.C.; "Crystal structure of bovine procarboxypeptidase A-S6 subunit III, a highly structured truncaed zymogen E."; EMBO J. 13:1763-1771(1994).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            TO HAVE A PROTERASE ACTIVITY. ITS LIKELY ENLICASE. DUES NOT SER PROCARBOXYPEPTIDASE A AGAINST DENATURATION IN THE ACIDIC ENVIRONMENT OF THE RUMINARY DUODENUM.

-I - SUBBNIT: HETEROTRIMER OF SUBBNIT III; CARBOXYPEPTIDASE A AND CHYMOTRYFSINGEN C.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Proproteinase E precursor (Procarboxypeptidase A complex component III) (Procarboxypeptidase A-56 subunit III) (PROCPA-S6 III).

Bost raurus (Bovine).

Eukaryota: Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria: Cetartiodactyla; Ruminantia; Pecora; Bovoidea; Bovinae; Bos.
134 SYL-----KGSSGGPLI.CPAGHAVGIFRAAVCTRCVAKAVDF-IPVESL 176
                                              266 QDVIQTDAAINPGNSGGPLLDSSGTLIGINTAIYSPSGASSGVGFSIPVDTV 317
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Venot N., Sciaky M., Puigserver A., Desnuelle P., Laurent G., "Amino acid sequence and disulfide bridges of subunit III, a defective endopeptidase present in the bovine pancreatic 6 S
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    [1]
SEQUENCE OF 1-25.
MEDLINE-91099520; PubMed-2269366;
Pascual R., Vendrell J., Aviles F.X., Bonicel J., Wicker C.,
```

```
Search completed: August 30, 2003, 19:13:43 Job time : 13.0138 secs
                                                                                                                                     EMBL: U05782; AAA74392.1; -.
PIR: S70401; S70401.
Interpro: IPR001507; Endoglin/CD105.
Pfan: PF00100; zona_pellucida; 1.
SMART; SM00241; ZP; 1.
                                                                                                                                                                                                                   44987 MW;
                                                                                                                                                                                    18
415
378
399
415
                                                                                                                                                                                                                        Query Match
Best Local Similarity
Thes 31; Conserva
                                                                                                                                                                                                                   415 AA;
                                               SEQUENCE FROM N.A.
                                    NCBI_TaxID=9986;
                ZPC.
                                                                                                                                                                                         CHAIN
DOMAIN
TRANSMEM
                                                                                                                                                                                                                    SEQUENCE
                                                                                                                                                                              NON_TER
SIGNAL
ö
                                                                                                                                                                                                                                                            q
                                                                                                                                                                                                                                                                       ŏ
                                                                                                                                                                                                                                                  10;
                                                                                                                                                                                                                                                                   || || || : | |:::|| |:: | |: ::| | CISTSRTYQVVLGEYDRSVLEGSEQVIPINAGDLFVHPLWNSNCVACGNDIALVKLSRSA 116
                                                                                                                                                                                                                                                                                        100 -CGSSDLYLVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAG-----HAV 152
                                                                                                                                                                                                                                                                                                                170 HCSQWDWWCITVKKIM--VCAGGDTR-----SGCNGDSGGPLNCPAADGSWQVHGV 218
                                                                                                                                                                                                                                                             99
                                                                                                                                                                                                                                                                                  -----SPKGPVI-----9MYTNVDKDLVGWPAPQG-SRSLTPCT--- 99
                                                                                                                                                                                                                                                            17 CIITSLTGR-----DKNQVEGEVQIVS-TAAQTFLATCINGVCWTVYHGAGTRTIA---
                                                                                                                                                                                                                                                  62; Gaps
                                                                                                                                                                                                                                       7.9%; Score 75; DB 1; Length 253; 22.7%; Pred. No. 4.9; 72; Indels ive 23; Mismatches 72; Indels
                                                                                                                                                                                                                             27337 MW; 24663724DBAE409C CRC54;
                                                                                                                                                                                                                                                                                                                                                                         (Rel. 33, Created)
(Rel. 33, Last sequence update)
(Rel. 33, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                415 AA.
                                                                                                                                                                                                                                                                                                                          GIFRAAVCTRGVAKAVDFIPVES 175
                                                                                                                                                                                                                                                                                                                                    219 TSFVSAFGCNTIKKPIVFTRVSA 241
                                                                                                                                                                                                                                                                                                                                                                PRT:
                                                                                                                                                                                                                                                  Conservative
                                                                                                                                                                                                                                                                                                                                                                STANDARD;
                                                                                   253 AA;
                                                                                                                                                                                                                                             Similarity
                                                                                        2P3_RABIT
P48B33;
01-FEB-1996 (
01-FEB-1996 (
01-FEB-1996 (
                                                                                                                                                                                                                                        Query Match
Best Local Simi
Matches 46;
                                                                                                                                                                                                                   HELIX
HELIX
SEQUENCE
                                                                                                                                                                                                                                                                                67
                                                                                                                                                                                                                                                                                                                          153
                                                  STRAND
STRAND
STRAND
TURN
STRAND
HELIX
TURN
STRAND
                                                                                                                                                                                                              STRAND
STRAND
                                                                                                                                                                                                                                                                                                                                                    RESULT 15
2P3_RABIT
ID 2P3_R
AC P4883
DT 01-FE
DT 01-FE
g
                                                                                                                                                                                                                                                                                          q
                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                              QQ
                                                                                                                                                                                                                                                                                                                          à
                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                             ò
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss institute of Bioinformatics and the EMBL outstation the European Bloinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ς.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              55 TVYHGAGTRIIASPKGP-VIQMYINVDKDLVGWPAPQGSRSLIPCTCGSSDLYLVTRHAD 113
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                327
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          "Cloning and characterization of zona pellucida genes and cDNAs from a variety of mammalian species: the ZPA, ZPB and ZPC gene families."; DNA Seq. 4:361-393(1994).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  114 VIPVRRRGDSRGSLLSPRPISYLKGSSGGPL--LCPAGHAVGIFRAAVCTRGVAKA 167
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Indels 15; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             20NA PELLUCIDA SPERM-BINDING PROTEIN EXTRACELLULAR (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                        TISSUE-Ovary;
MEDLINE-95143578; PubMed-7841460;
Harris J.D., Hibler D.W., Fontenot G.K., Hsu K.T., Yurewicz E.C.,
                                                                                                                                   Oryctolagus cuniculus (Rabbit).
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalla; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
2ona pellucida sperm-binding protein 3 precursor (2ona pellucida
glycoprotein 2P3) (Sperm receptor) (2ona pellucida protein C)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              PROSITE; PSOUG62; ZP_DOMAIN; 1.
Glycoprotein; Signal; Sulfation; Sperm; Receptor; Transmembrane;
Extracellular matrix; Multigene family.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           DB 1; Length 415;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  77396CF1BAA3F5CB CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CYTOPLASMIC (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   7.8%; Score 74.5; DB 1; 1 ilarity 26.7%; Pred. No. 9.5; Conservative 17; Mismatches 53;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      POTENTIAL
```

Н

00000000000000000000000000000000

hepatitis hepatitis

hepatitis hepatitis hepatitis hepatitis hepatitis hepatitis hepatitis hepatitis hepatitis

hepatitis

Run ĕ

```
2 APITAYAQQTRGLLGCIITSLTGRDKNQVEGEVQIVSTAAQTFLATCINGVCWTVYHGAG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            STRAIN-PL.1Y;
HOlland-Staley C.A., Kovari L.C., Golenberg E., Mayers D.L.;
Holland-Staley C.A., Kovari L.C., Golenberg E., Mayers D.L.;
Gonetic Diversity and response to IFN of the NS3 Protease Gene from
Clinical Strains of the Hopatitis C Virus.';
Submitted (APR-2001) to the EMBL/GenBank/DDBJ databases.
EMBL, ARK94560.1;
-1. InterPro; IPR004109; HCV_NS3:
Pfam: PF02907; HCV_NS3: 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                       no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ö
                                                                                    091779
0921779
0921772
0931772
0931772
0931773
0931773
0931773
0931773
0931773
0931774
0931774
0931774
0931774
                               091rr5
091rs4
091rt6
091rt9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  99.2%: Score 945; DB 12; Length 181; 99.4%; Pred. No. 1.9e-88; ive 1; Mismatches 0; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1 1
181 181
181 AA: 19130 MW; 85D91869299B7C35 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                           Last sequence update)
Last annotation update)
                                                                                                                                                                                                                                                                                                                                        ALIGNMENTS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Viruses; ssRNA positive-strand viruses, Hepacivirus
                                                                                                                                                                                                                                                                                                                                                                                                                  Created)
                              091RR5
091RR5
091RR4
091RR4
091RR7
091RR7
091RR7
091RR7
091RR7
091RR1
091RR1
091RR1
091RR2
091RR3
099RR3
09
                                                                                                                                                                                                                                                                                                                                                                                            PRT;
                                                                                                                                                                                                                                                                                                                                                                                          O91RR8 PRELIMINARY;
O91CBC-2001 (TEMBLE) 19, Ci
01-DEC-2001 (TEMBLE) 19, Li
01-DEC-2001 (TEMBLE) 19, Li
N33 protease (Fragment).
Hepatitis C virus.
 Conservative
 Best Local Similarity
Matches 180; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           NCBI_TaxID-11103;
62
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Protease
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Query Match
                                                                                                                                                                                                                                                                                                                                                                     RESULT 1
091RR8
 ag
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      091rr8 hepatitis c
091rt5 hepatitis c
091fe5 hepatitis c
091rs hepatitis c
091rs hepatitis c
091rq8 hepatitis c
091rq9 hepatitis c
091rs hepatitis c
                                                                      ; Search time 34.7298 Seconds (without alignments) 1352.314 Million cell updates/sec
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.
                                                                                                                                       1 MAPITAYAQQTRGLLGCIII.......GVAKAVDFIPVESLETTMRS 182
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Description
        5.1.6
Compugen Ltd.
                                                                                                                                                                                                               Total number of hits satisfying chosen parameters:
                                                                                                                                                                                          830525 seqs, 258052604 residues
        GenCore version
Copyright (c) :993 - 2003
                                                                                                                                                                                                                                                                                             summaries
                                                   - protein search, using sw model
                                                                       August 30, 2003, 19:00:22
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        091RR8
091RT5
081RT5
091RE3
091RR1
091RR1
091RR6
091RR6
091RR8
091RR8
091RR8
091RR8
                                                                                                                                                                       Gapop 10.0 , Gapext 0.5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                    sp_unclassified:*
                                                                                                                                                                                                                                                                                                                                                                                                 sp_organelle:*
sp_phage:*
sp_lant:*
sp_rodent:*
sp_vivus:*
sp_virus:*
                                                                                                                                                                                                                                                                                                                                                                      sp_invertebrate:*
                                                                                                                                                                                                                                                                                 Maximum Match 100%
Listing first 45 su
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          sp_bacteriap:*
                                                                                                                                                                                                                                                                                                                SPTREMBL_23:*
: sp_archea:*
: sp_bacteria:*
                                                                                                                                                                                                                                     seq length: 0 seq length: 20000000000
                                                                                                                                                                                                                                                                    Post-processing: Minimum Match 0%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          sp_archeap:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   53
                                                                                                                                                                                                                                                                                                                                                                                 sp_mammal:*
                                                                                                                  US-09-965-594-1
                                                                                                                                                                                                                                                                                                                                                Sp_fungi:*
Sp_human:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Query
Match Length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         sp_mhc:*
                                                                                                                                                            BLOSUM62
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ဝဝထထထထက်က်ကဲလဲလ
                                                                                                                                                                                                                                                                                                                                                                                                                        10:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Perfect score:
                                                                                                                                                            Scoring table:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Score
                                                                                                                                                                                                                                    Minimum DB s
Maximum DB s
                                                     protein
                                                                                                                                                                                            Searched:
                                                                                                                                        Sequence:
                                                                                                                                                                                                                                                                                                                Database
                                                                        on:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          6543210987654321
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Result
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Š
```

121 120

09

61

Gaps

```
(MAY-1990)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    11
RS 757
                                                                                                                        InterPro;
InterPro;
InterPro;
InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             NON_TER
NON_TER
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       756
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Q9IFES
Q9IFES;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         RESULT 4
      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Ω
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  å
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         οχ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           9
DSRGSLLSPRPISYLKGSSCGPLLCPACHAVGIFRAAVCTRGVAKAVDFIPVESLETIMR 181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  2 APITAYAQQTRGLLGCIITSLTGRDKNQVEGEVQIVSTAAQTFLATCINGVCWTVYHGAG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1 APITAYAQQTRCLIGCIIISLIGRBKNQVRCEVQIVSTAAGTFIATCINGVCWTVYHGAG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          62 TRIIASPKGPVIQMYINVDKDLVGWPAPQGSRSLTPCTCGSSDLYLVIRHADVIPVRRRG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          STRAIN-Pt.4;
Holland-Stelley C.A., Kovari L.C., Golenberg E., Mayers D.L.;
Holland-Stelley C.A., Kovari L.C., Golenberg E., Mayers D.L.;
Genetic Diversity and response to IFN of the NS3 Protease Gene from
Clinical Strains of the Hepatilis C Virus.*;
Submitted (APR-2001) to the EMBL/GenBank/DDBJ databases.
EMBL, AF869218; AAK54543.1;
InterPro; IPR004109; HCV.NS3.
Pfam; PF02907; HCV_NS3; I.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Genome polyprotein (Fragment).
Hepatitis C virus.
Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
Hepacivirus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ;
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Score 945; DB i2; Length 181;
Pred, No. 1.9e-88;
1; Mismatches 0; Indels (
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                SEQUENCE FROM N.A. Choo Q.-L., Richman K., Han J.; "The nucleotide sequence of the Hepatitis C viral genome.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  181 181
181 AA; 19130 MW; 85D9186929987C35 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                               01-DEC-2001 (TrEMBLrel. 19, Created)
01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
01-DAR-2003 (TrEMBLrel. 23, Last annotation update)
NS3 procease (Fragment).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Q81756 PRELIMINARY; PRT; 2436 AA. Q81756; Ol.NOV-1996 (TrEMBLrel. 01, Created) 01-NOV-1996 (TrEMBLrel. 01, Last sequence update) 01-NOV-2903 (TrEMBLrel. 23, Last annotation update)
                                                                                                                                                                                                                                                                                                                                           181 AA.
                                                                                                                                                                                                                                                                                                                                           PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             99.2%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           tches 180; Conservative
                                                                                                                                                                                                                                                                                                                                           PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         NCBI_TaxID-11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         NCBI_TaxID-11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        S 182
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               S 181
                                                                                                                           182 S 182
                                                                                                                                                                                  S 181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Hepacivirus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     NON_TER
NON_TER
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Protease.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        182
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  181
122
                                                                                                                                                                                      181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Local
                                                                                                                                                                                                                                                                                                                                           Q91RT5
Q91RT5;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             RESULT 3
08176
08176
08176
08170
08170
08170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
09170
091
                                                                                                                                                                                                                                                                             RESULT 2
(91875)
10 (91875)
10 (91876)
10 (91876)
10 (91876)
10 (91876)
10 (91876)
10 (91876)
10 (91876)
10 (91876)
10 (91876)
10 (91876)
10 (91876)
10 (91876)
10 (91876)
10 (91876)
10 (91876)
10 (91876)
10 (91876)
10 (91876)
10 (91876)
10 (91876)
10 (91876)
10 (91876)
10 (91876)
10 (91876)
10 (91876)
10 (91876)
10 (91876)
10 (91876)
10 (91876)
10 (91876)
10 (91876)
10 (91876)
10 (91876)
10 (91876)
10 (91876)
10 (91876)
10 (91876)
10 (91876)
10 (91876)
10 (91876)
10 (91876)
10 (91876)
10 (91876)
10 (91876)
10 (91876)
10 (91876)
10 (91876)
10 (91876)
10 (91876)
10 (91876)
10 (91876)
10 (91876)
10 (91876)
10 (91876)
10 (91876)
10 (91876)
10 (91876)
10 (91876)
10 (91876)
10 (91876)
10 (91876)
10 (91876)
10 (91876)
10 (91876)
10 (91876)
10 (91876)
10 (91876)
10 (91876)
10 (91876)
10 (91876)
10 (91876)
10 (91876)
10 (91876)
10 (91876)
10 (91876)
10 (91876)
10 (91876)
10 (91876)
10 (91876)
10 (91876)
10 (91876)
10 (91876)
10 (91876)
10 (91876)
10 (91876)
10 (91876)
10 (91876)
10 (91876)
10 (91876)
10 (91876)
10 (91876)
10 (91876)
10 (91876)
10 (91876)
10 (91876)
10 (91876)
10 (91876)
10 (91876)
10 (91876)
10 (91876)
10 (91876)
10 (91876)
10 (91876)
10 (91876)
10 (91876)
10 (91876)
10 (91876)
10 (91876)
10 (91876)
10 (91876)
10 (91876)
10 (91876)
10 (91876)
10 (91876)
10 (91876)
10 (91876)
10 (91876)
10 (91876)
10 (91876)
10 (91876)
10 (91876)
10 (91876)
10 (91876)
10 (91876)
10 (91876)
10 (91876)
10 (91876)
10 (91876)
10 (91876)
10 (91876)
10 (91876)
10 (91876)
10 (91876)
10 (91876)
10 (91876)
10 (91876)
10 (91876)
10 (91876)
10 (91876)
10 (91876)
10 (91876)
10 (91876)
10 (91876)
10 (91876)
10 (91876)
10 (91876)
10 (91876)
10 (91876)
10 (91876)
10 (91876)
10 (91876)
10 (91876)
10 (91876)
10 (91876)
10 (91876)
10 (91876)
10 (91876)
10 (91876)
10 (91876)
10 (91876)
10 (91876)
10 (91876)
10 (91876)
10 (91876)
10 (91876)
10 (91876)
10 (91876)
10 (91876)
10 (91876)
10 (91876)
10 (91876)
10 (91876)
10 (91876)
10 (91876)
10 (91876)
10 (91876)
10 (91876)
10 (91876)
10 (91876)
10 (91876)
10 (91876)
10 (91876)
10 (91876)
10 (9
                                                             a
                                                                                                                                                                               οp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Ω
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                           à
ö
```

```
61 GTRIIASPKGPVIQMYTNVDKDLVGWPAPQGSRSLIPCTCGSSDLYLVTRHADVIPVRRR 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        GDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVCTRGVAKAVDFIPVESLETTM 180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1 MAPITAYAQQIRGLIGCIIISLIGRDKNQVEGEVQIVSTAAQIFLATCINGVCWIVYHGA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          MEDINE-21262212; PubMed-11369872;
Lanford R.E., Lee H., Chavez D., Guerra B., Brasky K.M.;
"Infectious cDNA clone of the hepatitis C virus genotype 1 prototype
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  0; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ATP-binding; Coat profein; Envelope protein; Glycoprotein; Hellcase;
Hydrolase; Nonstructural protein; Polyprotein;
RNA-directed RNA polymerase; Transferase; Transmembrane.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Genome polyprotein.
Hepatitis C virus.
Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
Hepatuirus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   99.0%; Score 943; DB 12; Length 2436; 98.4%; Pred. No. 7.1e-87; tive 3; Mismatches 0; Indels 0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           264734 MW; D7B9872900BE3125 CRC64;
EMBL/GenBank/DDBJ databases.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            01-0cT-2000 (TrEMBLrel. 15, Created)
01-0cT-2000 (TrEMBLrel. 15, Last sequence update)
01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       PRT; 3011 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                PROSITE; PS50507; RDRP_POSITIVE; 1.
PROSITE; PS50521; RDRP_VIRAL; 1.
                                                                                                                                                          InterPro; IPR001490; HCV_NG4b.
InterPro; IPR002868; HCV_NG5a.
InterPro; IPR002166; HCV_RGRP.
InterPro; IPR001650; Helicase_C.
InterPro; IPR007095; RNA_pol_DS_PS.
InterPro; IPR007094; RNA_pol_DS_PS.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          seguence.";
J. Gen. Virol. 82:1291-1297(2001).
                                                                                                                                                                                                                                                                                      Pram; PF01560; HCV_NS1; I.Pfam; PF01560; HCV_NS1; I.Pfam; PF012907; HCV_NS3; I.Pfam; PF01006; HCV_NS4a; I.Pfam; PF01006; HCV_NS4b; I.Pfam; PF010506; HCV_NS4b; I.Pfam; PF00271; helicase_C; I.Pfam; PF00271; helicase_C; I.Probom; PF00988; Viral_RGRP; I.Probom; PR00988; Viral_RGRP; I.SMART; SMO0487; DEXDC; I.
                                                                                                                                           IPR000745; HCV_NS4a
  the
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Best Local Similarity 98.4
Matches 179; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            2436
                                                                                                     IPR002518;
IPR004109;
                                                             InterPro; IPR001410;
                                                                                 IPR002531;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          2436 24;
2436 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (1)
SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              NCBI_TaxID=11103;
```

ö

635

 $\sim$ 

```
Best Local Similarity 98.9
Matches 179; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Best Local Similarity 98.9
Matches 179; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                NS3 protease (Fragment).
Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SEQUENCE FROM N.A. STRAIN-Pt.K;
                                                                                                                                                        SEQUENCE FROM N.A.
                                                      NCBI_TaxID-111.03;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      NCBI_TaxID-11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             182 S 182
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     s 181
              Hepacivirus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Protease.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               NON_TER
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          NON_TER
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Protease
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   NON_TER
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             091RS1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          RESULT 6
Q91RS1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             HD DDT TO DDT TO
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
              OCC OX REPRESENT REPRESENT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           \delta
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               d
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Qγ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     61 GTRTIASPKGPVIQMYTNVDKDLVGWPAPQGSRSLTPCTCGSSDLYLVTRHADVIPVRRR 126
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    121 GDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVCTRGVAKAVDFIPVESLETTM 180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 09
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           | InterPro; | IFR002166; | HCV_ACADAGO
| InterPro; | IFR002166; | HCV_ACADAGO
| InterPro; | IFR001050; | HCV_ACADAGO
| InterPro; | IFR001050; | HCV_ACADAGO
| InterPro; | IFR007094; | RNA_POI_BS_ES.
| InterPro; | IFR007094; | RNA_POI_BS_IS.
| InterPro; | IFR007094; | RNA_POI_BS_IS.
| InterPro; | IFR007094; | INTERPRO; | INTERPRO;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1 MAPITAYAQQTRGLLGCIITSLTGRDKNQVEGEVQIVSTAAQTFLATCINGVCWTVYHGA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   01-DEC-2001 (TrEMBLrel. 19, Created)
01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
NG3 protease (Fragment).
Hepatitis C virus.
Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY LIPOPROTEIN BYNUELOPE. CONSISTS OF TWO PROFINS: PROFIEIN M AND GIYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C. AND MRNA (BY SIMILARITY).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ċ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Score 943; DB 12; Length 3011; Pred. No. 9.2e-87;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   0; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Hydrolase; Nonstructural protein; Polyprotein;
RNA-directed RNA polymerase; Transferase; Transmembrane.
SEQUENCE 3011 AA; 327124 WW; 2489CE74AC864E58 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Mismatches
                                                                                                                                                                                                                                                                                        InterPro; IPR000345; CytC_heme_bind.
InterPro; IPR001410; DEAD.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              PRT;
                                                                                                                                                                                                                                                                                                                                                                            InterPro: IPR002522; HCV_capsid.
InterPro: IPR002521; HCV_capsid.
InterPro: IPR002519; HCV_core.
InterPro: IPR002519; HCV_NS1.
InterPro: IPR004109; HCV_NS2.
InterPro: IPR004109; HCV_NS3.
InterPro: IPR000745; HCV_NS3.
InterPro: IPR001490; HCV_NS4.
InterPro: IPR001490; HCV_NS4.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   3;
                                                                                                                                                                                                EMBL; AF271632; AAF81759.1; -.
HSSP; P27958; 1A1V.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            99.08;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Matches 179; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       RS 1207
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    RS 182
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Q91RR3;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Q91RR3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 RESULT 5
    <del>:</del>.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Q91RR3
1D Q9
AC Q9
DT Q1
DT Q1
DT Q1
OS HE
    SON WEEK WEEK WEEK WAS A STREET OF A STREE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          op
O
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
```

```
ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                TRTIASPKGPVIQMYINVDKDLVGWPAPQGSRSLIPCTCGSSDLYLVIRHADVIPVRRG 121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    DSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVCTRGVAKAVDFIPVESLETTMR 181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1 APITAYAQQTRGLLGCIITSLTGRDKNQVEGEVQIVSTAAQTFLATCINGVCWTVYHGAG 60
                                                                                                                                                                                                                                                                                                                                                                                                                                                              9
                                                                                                                                                                                                                                                                                                                                                                                                               2 APITAYAQQIRGLLGCIITSLIGRDKNQVEGEVQIVSTAAQIFLAICINGVCWTVYHGAG 61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          2 APITAYAQQTRGLLGCIITSLTGRDKNQVEGEVQIVSTAAQTFLATCINGVCWTVYHGAG
                                                                                                                                                                                                                                                                                                                                                                                                                                             Holland-Staley C.A., Kovari L.C., Golenberg E., Mayers D.L.;
"Genetic Diversity and response to IFN of the NS3 Protease Gene from Clinical Strains of the Hepatitis C Virus.";
Submitted (ARR-2001) to the EMBL/GenBank/DDBJ databases.
EMBL: AFS69232: AAK54557.1;
InterPro: IPR004109: HCV_NS3.
Pfam; PF02907; HCV_NS3; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Gaps
STRAIN-Pt.(B):
Holland-Staley C.A., Kovari L.C., Golenberg E., Mayers D.L.;
Holland-Staley C.A., Kovari L.C., Golenberg E., Mayers D.L.;
Genetic Diversity and response to IFN of the NS3 Protease Gene from
Clinical Strains of the Hepatitis C Virus.";
Submitted (APR-2001) to the EMBL/GenBank/DDBJ databases.
EMBL, AFS69240; AAK54565.1;
InterPro; IPR004109; HCV NS3.
Pfam; PF02907; HCV_NS3; I.
                                                                                                                                                                                                                                                                                                                                                                    Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
Hepacivirus.
                                                                                                                                                                                                                                                                                                                                                                    ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ô
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Length 181;
                                                                                                                                                                                                                                                                                                                  Length 181;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Indels
                                                                                                                                                                                                                                                                                                                                                                    Indels
                                                                                                                                                                                                                                            181 181
181 AA; 19115 MW; 5D85F88AD7AC1A11 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  181 181
181 AA; 19114 MW; ABB90B5B3ABA4E26 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Created)
Last sequence update)
Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                  .;
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        98.8%; Score 942; DB 12;
98.9%; Pred. No. 3.8e-88;
Live 2; Mismatches 0;
                                                                                                                                                                                                                                                                                                                  Score 942; DB 12;
Pred. No. 3.8e-88;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      181 AA
                                                                                                                                                                                                                                                                                                                                                                  2; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 PRT;
                                                                                                                                                                                                                                                                                                                  98.8%;
98.9%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                01-DEC-2001 (TrEMBLrel. 19, 01-DEC-2001 (TrEMBLrel. 19, 01-MAR-2003 (TrEMBLrel. 23,
```

```
PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SEQUENCE FROM N.A.
 [1]
SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                             NCBI_TaxID-11103;
                                                                                                                                                                                                                                                                                                      182 S 182
                                                                                                                                                                                                                                                                                                                        S 181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          STRAIN-Pt. 3T
                                                                                                                                                                                                                                                                                                                                                                                                                                                     Hepacivirus
                                                                                                        NON_TER
NON_TER
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         NON_TER
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   62
                                                                                                                                                                                                                                                                                                                        181
                                                                                               Protease.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Protease
                                                                                                                                               Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               NON_TER
                                                                                                                                                                                                                                                                                                                                                                        Q91RR6
                                                                                                                                                                                                                                                                                                                                                    RESULT 9
091RR6
  ò
                                                                                                                                                                                                      qq
                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                           a
                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                       d
                                                                                                                                                                                                                                                                                                                                                                                  ô
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ô
                                                                                                                                                                                                                                                                                                                                                                                                               :.
ن
                                                                                                                                                                                                                                                                                                                                                                                                                                                                               121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            DSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVCTRGVAKAVDFIPVESLETTMR 181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       61 TRIIASPKGPVIQMYINVDKDLVGMPAPQGARSLIPCTCGSSDLYLVTRHADVTPVRRG 120
                                                 TRIIASPKGPVIQMYINVDKDLVGWPAPQGSRSLTPCTCGSSDLYLVIRHADVIPVRRRG 121
                                       122 DSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVCTRGVAKAVDFIPVESLETTMR 181
                                                                                                                                                                                                                                                                                                                                                                                                                                                   09
                                                                                                                                                                                                                                                                                                                                                                                                                                          2 APITAYAQQTRGLLGCIITSLTGRDKNQVEGEVOIVSIAAQTFLATCINGVCWTVYHGAG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                       TRTIASPKGPVIQMYTNVDKDLVGWPAPQGSRSLTPCTCGSSDLYLVTRHADVIPVRRRG
                                                                                                                                                                                                                                                                      Holland-Staley C.A., Kovari L.C., Golonberg E., Mayers D.J.;
"Genedic Diversity and response to IrN of the NS3 Protease Gene from
Clinical Strains of the Hepatitis C Virus.";
Submitted (APR-2001) to the EMBL/GenBank/DDBJ databases.
EMBL, AF856245; AR854570.1;
InterPro; IPR004109; HCV_NS3.
                                                                                                                                                                                                           Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
Hepacivirus.
NCBI_TaxID=11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                                              ;
                                                                                                                                                                                                                                                                                                                                                                                         98.9%; Score 942; DB 12; Length 181; 98.9%; Pred. No. 3.8e-88; 1ve 2; Mismatches 0; Indeis
                                                                                                                                                                                                                                                                                                                                                             181 181
181 AA: 19144 MW; COC91F1E2EEB0B32 CRC64;
                                                                                                                                                               01-DEC-2001 (TrEMBLrel. 19, Created)
01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     01-DEC-2001 (TIEMBLrel. 19, Created)
01-DEC-2001 (TIEMBLrel. 19, Last sequence update)
01-MAR-2003 (TIEMBLrel. 23, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     181 AA
                                                                                                                                             181 AA
                                                                                                                                              PRT;
                                                                                                                                                                                                                                                                                                                                                                                                              Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   NS3 protease (Fragment). Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     PRELIMINARY;
                                                                                                                                                                                             NS3 protease (Fragment). Hepatitis C virus.
                                                                                                                                             PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                   Best Local Similarity
Matches 179; Conserv
                                                                                                                                                                                                                                                     SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         NCBI_TaxID=11103;
                                                                                          s 181
                                                                           182 S 182
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  S 182
                                                                                                                                                                                                                                                               STRAIN=Pt.5Z;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Hepacivirus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   S
                                                                                                                                                                                                                                                                                                                                         Protease.
                                                                                                                                                                                                                                                                                                                                                             NON TER
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                      62
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           122
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  182
                                                                                             181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   181
                                                                                                                                                                                                                                                                                                                                                                                           Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              091RT1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    091RT1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                RESULT 8
091RT1
1D 091RR
AC 091RR
DT 01-D
DT 01-D
DT 01-D
OC NITU
OC Hepa
                                                                                                                                  õ
                                                       g
                                                                           ò
                                                                                            a
                                                                                                                                                                                                                                                                                                                                                                                                                                 ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ô
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
ò
                   a
```

```
ö
                                                                                                                                                                                                                                                                                                                                            ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               61 TRIIASPKGPVIQMYINVDKDLVGWPAPQGARSLIPCICGSSDLYLVIRHADVIPVRRG 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    62 TRTIASPKGPVIQMYTNVDKDLVGWPAPQGSRSLTPCTCGSSDLYLVTRHADVIPVRRG 121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    2 APITAYAQQTRGLLGCIITSLTGRDKNQVEGEVQIVSTAAQTFLATCINGVCWTVYHGAG 61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              61
                                                                                                                                                                                                                                                                                                                                                                                                                    2 APITAYAQQTRGLLGCIITSLTGRDKNQVEGEVQIVSTAAQTFLATCINGVCWTVYHGAG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    122 DSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVCTRGVAKAVDFIPVESLETTMR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Gaps
Holland-Staley C.A., Kovari L.C., Golenberg E., Mayers D.L.;
"Genetic Diversity and response to Irw of the NS3 Protease Gene from
"Clinical Strains of the Hepatitis C Virus.";
Submitted (APR-2001) to the EMBL/GenBank/DDBJ databases.
EMBL, AF865221, AAR5447.1:
Interpro, IPR004109; HCV NS3.
                                                                                                                                                                                                                                                                                                                                               Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ö
                                                                                                                                                                                                                                                                                                                                            ;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Query Match 98.6%; Score 940; DB 12; Length 181; Best Local Similarity 98.3%; Pred. No. 6e-88; Matches 178; Conservative 3; Mismatches 0; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Holland-Staley C.A., Kovari L.C., Golenberg E., Mayers D.L., Genetic Diversity and response to IFN of the NS3 Protease G Clinical Strains of the Hepatitis C Virus."; Submitted (APR-2001) to the EMBL/GenBank/DDBJ databases. EMBL, AFS89237. AAR54562.1; InterPro; IPR004109; HCV.NS3.
                                                                                                                                                                                                                                                                                           Length 181;
                                                                                                                                                                                                                                                                                                                                          Indels
                                                                                                                                                                                              1 1
181 181
181 AA; ABB90B5B3ABA4E26 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       181 AA; 19101 MW; 614ADABBOF33CCAF CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Last sequence update)
Last annotation update)
                                                                                                                                                                                                                                                                                                                                            .;
0
                                                                                                                                                                                                                                                                                      98.9%; Score 942; DB 12;
98.9%; Pred. No. 3.8e-88;
tive 2; Mismatches 0;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Created)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           091RR6;
01-DEC-2001 (TrEMBLrel. 19, C.
01-DEC-2001 (TrEMBLrel. 19, L.
01-MAR-2003 (TrEMBLrel. 23, L.
NS3 protease (Fragment)
Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                  Best Local Similarity 98.93
Matches 179; Conservative
```

S

유 à g à Q

```
122 DSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVCTRGVAKAVDFIPVESLETTMR 181
                                                                                                                                                                                                                                                                                                                                                                                                   STRAIN-HC-JI; MEDLINE-9204440; PubMed-1658196; MEDLINE-9204440; PubMed-1658196; Okamoto H., Machida A., Okamoto H., Mayumi M.; Miyakawa Y., Mayumi M.; Mucleotide sequences of the genomic RNA of hepatitis C virus isolated from a human carrier. comparison with reported isolates for conserved and divergent regions. "). Gen. Virol. 72:2697-2704(1991).
                                                                                                                                                                                                                                                                         62 TRTIASPKGPVIQMYTNVDKDLVGWPAPQGSRSLTPCTCGSSDLYLVTRHADVIPVRRRG
                                                                                                                                                                                                                                                               2 APITAYAQOTRGLLGCIITSLTGRDKNQVEGEVQIVSTAAQTFLATCINGVCWTVYHGAG
                              "Genetic Diversity and response to IFN of the NS3 Protease Gene from Clinical Strains of the Hepatitis C Virus.";
Submitted (ASR-2001) to the EMBL/GenBank/DDBJ databases.
EMBL: AFS69230; AAK54555.1;
InterPro: IPR004109; HCV_NS3.
Pfam; PF02907; HCV_NS3; 1.
                                                                                                                                                                                                                                 Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
Hepacivirus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        MEDILINE-33117120; PubMed-1335573;
Okamoto H., Kanai N., Mishiro S.;
"Full-length nucleotide sequence of a Japanese hepatitis C virus isolate (HC-11) with high homology to USA isolates.";
Nucleic Acids Res. 20:6410-6410(1992).
                                                                                                                                                                                                                                 0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             "The 5'-terminal sequence of the hepatitis C virus genome.";
Jpn. J. Exp. Med. 60:167-177(1990).
                   Holland-Staley C.A., Kovari L.C., Golenberg E., Mayers D.L.;
                                                                                                                                                                                                      Length 181;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Okamoto H., Okada S., Sugiyama Y., Yotsumoto S., Tanaka T.,
Yoshizawa H.;
                                                                                                                                                                                                                               0; Indels
                                                                                                                                                      181
19132 MW; 0BB90B5F3AB95250 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Last sequence update)
Last annotation update)
                                                                                                                                                                                                   98.5%; Score 939; DB 12; 98.3%; Pred. No. 7.6e-88; iive 3; Mismatches 0;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     3011 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Created)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  STRAIN-HC-JI;
MEDLINE-91013116; PubMed-2170712;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                01-NOV-1996 (TrEMBLrel. 01,
01-NOV-1996 (TrEMBLrel. 01,
01-MAR-2003 (TrEMBLrel. 23,
                                                                                                                                                                                                                 Local Similarity 98.3 ies 178; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     PRELIMINARY;
                                                                                                                                       1
181
181 AA; 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Genome polyprotein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         NCBI_TaxID-11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                              182 S 182
                                                                                                                                                                                                                                                                                                                                                                                                                                                                           S 181
   STRAIN-Pt.24;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Hepatitis C
                                                                                                                          Protease.
                                                                                                                                        NON_TER
NON_TER
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                            181
                                                                                                                                                                                                    Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Q03463
Q03463;
                                                                                                                                                                                                                                 Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       RESULT 12
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     003463
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   셤
                                                                                                                                                                                                                                                                                                                                                   ΩΩ
                                                                                                                                                                                                                                                               á
                                                                                                                                                                                                                                                                                                                        à
                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                              qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                      셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   121 DSRGSILLSPRPISYLKGSSGGPLICPAGHAVGIPRAAVCTRGVAKAVDFIPVENLETTMR 180
                                            DSRGSLI.SPRPISYLKGSSCGPLLCPAGHAVGIFRAAVCTRGVAKAVDFIPVESI.ETTMR 181
61 TRIIASPKGPVIQMYINVDQDLVGWPAPQGARSLIPCTCGSSDLYLVTRHADV1PVRRRG 120
                               DSRGSLLSPRP1SY1,KGSSGGP1,LCPAGHAVGIFRAAVCIRGVAKAVDFIPVESLFTTMR 181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           2 APITAYAQQTRGLLGCIITSLTGRDKNQVEGEVQIVSTAAQTFLATCINGVCWTVYHGAG 61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                from
                                                                                                                                                                                                                                                                                                         Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                             ;
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                DB 12; Length 181;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  181 AA: 19131 MW; 8BD7FC2769DBD635 CRC64;
                                                                                                                                                                                                                            01-DEC-2001 (TrEMBLrel. 19, Created)
01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Last sequence update)
Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               98.6%; Score 940; DB 12;
98.9%; Pred. No. 6e-88;
Live 2; Mismatches 0;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    181 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Created)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 091RS3:
01-DEC-2001 (TEMBLEEL: 19,
01-DEC-2001 (TEMBLEEL: 19,
01-MAR-2003 (TEMBLEEL: 23,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Best Local Similarity 98.9
Matches 179; Conservative
                                                                                                                                                                                                                                                                         NS3 protease (Fragment).
Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            NS3 protease (Fragment). Hepatitis C virus.
                                                                                                                                                                                                   PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        [1]
SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                   SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                    NCBI_TaxID=11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         NCBI_TaxID=11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       181
                                                                                                                       S 181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   s 181
                                                                                          S 182
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           182 S 182
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Hepacivirus
                                                                                                                                                                                                                                                                                                                       Hepacivirus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Protease.
                                                            121
                                                                                          182
                                                                                                                     181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     NON_TER
```

ò g ö 8 ò Ω ò **091RS3** 

RESULT 11
091RS
AC 091RS
AC 091RS
DT 01-DE
DT 01-MB
DE NS3 P
OS Hepat
OC Hepat
OC Hepac
OX NCBI-

121

RESULT 10
791RS
AC Q91RS
AC Q91RS
DT 01-DE
DT 01-ME
DT 01

```
NON_TER
NON_TER
SEQUENCE
                                                                                                                                                                                                                                                                                                                        Protease.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Protease.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         181
                                                                                                                                                                                                                                                                                                                                                                                         Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Q91RS8
Q91RS8;
                                         091RT4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 RESULT 14
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 091RS8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1D
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ö
                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ő
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      GTRTIASPKGPVIQMYTNVDKDLVGWPAPQGSRSLTPCTCGSSDLYLVTRHADVIPVRRR 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           121 GDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVCTRGVAKAVDFIPVESLETTM 180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1 MAPITAYAQQTRGLLGCIITSLTGRDKNQVEGEVQIVSTAAQIFLATCINGVCWTVYHGA
                                                                                                                                 **Characterization and mapping of a B-cell immunogenic domain in hepatitis C virus E2 glycoprotein using a yeast peptide library."; Virology 200:246-255(1994).
-!- SUBGUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS: PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MRNA (BY SIMILARITY).

EMBL; D10749; BAA01582.1; -.
HSSP; P27958; IHEI.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     PROSITE: PSS6507; RDRP_105111VE; 1.
PROSITE: PSS6521; RDRP_VIRAL: 1.
PROSITE: PSS6521; RDRP_VIRAL: 1.
PROSITE: PSS6521; RDRP_VIRAL: 1.
Hydrolase: Nonstructural protein; Polyprotein;
RNA-directed RNA polymerase: Transferase: Transmembrane.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  0;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Query Match 98.5%; Score 939; DB 12; Length 3011; Best Local Similarity 97.8%; Pred. No. 2.4e-86; Matches 178; Conservative 3; Mismatches 1; Indels 0;
                                                                                                           Mink M., Benichou S., Madaule P., Tiollais P., Prince A.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           97E9052C0250463B CRC64;
                                       to the EMBL/GenBank/DDBJ databases
                                                                                                                                                                                                                                                       InterPro; IPR001410; DEAD.
InterPro; IPR001410; DEAD.
InterPro; IPR00252; HCV_capsid.
InterPro; IPR00251; HCV_core.
InterPro; IPR00251; HCV_core.
InterPro; IPR00251; HCV_NS1.
InterPro; IPR00419; HCV_NS2.
InterPro; IPR00419; HCV_NS3.
InterPro; IPR00149; HCV_NS3.
InterPro; IPR001490; HCV_NS4.
InterPro; IPR001490; HCV_NS4b.
InterPro; IPR00166; HCV_NS4b.
InterPro; IPR001669; HCV_NS5a.
InterPro; IPR001669; HCV_NS4b.
InterPro; IPR001095; RNA_pol_DS_PS.
                                                                                               MEDLINE-94174722; PubMed-7510436;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             327112 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                       Pram; PPO1515.
Pram; PPO1515.
Pram; PPO1542; HCV_copesid 1.
Pram; PPO1560; HCV_core; 1.
Pram; PPO1539; HCV_NS1; 1.
Pram; PPO1509; HCV_NS2; 1.
Pram; PPO1000; HCV_NS2; 1.
Pram; PPO1000; HCV_NS3; 1.
Pram; PPO1000; HCV_NS4a; 1.
Pram; PPO1001; HCV_NS4a; 1.
Pram; PPO1001; HCV_NS5a; 1.
Pram; PPO1001; HCV_NS5a; 1.
Pram; PPO1009; VITAL_RARP; 1.
ProDom; PD0196662; HCV_NS1; 1.
ProDom; PD0196662; HCV_NS1; 1.
                            Okamoto H.;
Submitted (DEC-1992)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             3011 AA;
                                                                   SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             RS 1207
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 RS 182
                STRAIN-HC-J1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1206
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Query Match
g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 οχ
```

```
61 TKTIASPKGPVIQMYTHVDQDLVGWPAPQGARSLTPCTCGSSDLYLVTRHADVIPVRRG 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        62 TRTIASPKGPVIQMYINVDKDLVGWPAPQGSRSLIPCTCGSSDLYLVTRHADVIPVRRRG 121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 122 DSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVCTRGVAKAVDFIPVESLETTMR 181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        2 APITAYAQQTRGLLGCIITSLTGRDKNQVEGEVQIVSTAAQTFLATCINGVCWTVYHGAG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Holland-Staley C.A., Kovari L.C., Golenberg E., Mayers D.L.;
"Genetic Diversity and response to IFN of the NS3 Protease Gene from Clinical Strains of the Hepatitis C Virus.";
Submitted (APR-2001) to the EMBL/GenBank/DDBJ databases.
EMBL; AS369225; AAK54550.1;
Interpro: IPR004109; HCV_NS3;
Pfam; PF02907; HCV_NS3; I.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Gaps
                                                                                                                                                                                                                                                                                   Holland-Staley C.A., Kovari L.C., Golenberg E., Mayers D.L.;
"Genetic Diversity and response to IFN of the NS3 Protease Gene from
Clinical Strains of the Hepatitis C Virus.";
Submitted (APR-2001) to the EMBL/GenBank/DDBJ databases.
EMBL, AF862219, AAR545441;
InterPro: IPR004109; HCV_NS3:
Pfam; PF02907; HCV_NS3:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
Hepacivirus.
                                                                                                                                            Viruses; ssRNA positive-strand viruses, no DNA stage; Plaviviridae;
Hepacivirus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ó:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           98.4%; Score 938; DB 12; Length 181; 98.3%; Pred. No. 9.6e-88;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1 1
181 181
181 AA; 19059 MW; 1E53C47AE8B7E5C9 CRC64;
                                        01-DEC-2001 (TrEMBLrel. 19, Created)
01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
01-DEC-2001 (TrEMBLrel. 23, Last annotation update)
NS3 protease (Fragment).
Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Last sequence update)
Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              181 AA
181 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Pred. No. 9.6e
3; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       01-DEC-2001 (TrEMBLrel. 19, Created)
01-DEC-2001 (TrEMBLrel. 19, Last sequ
01-MAR-2003 (TrEMBLrel. 23, Last anno
PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              PRELIMINARY;
PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         NS3 protease (Fragment)
Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Best Local Similarity
Matches 178; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SEQUENCE FROM N.A.
                                                                                                                                                                                                                                           SEQUENCE FROM N.A. STRAIN-Pt.23;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      NCBI_TaxID=11103;
                                                                                                                                                                                                NCBI_TaxID-11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      STRAIN-Pt. 176;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         182 S 182
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 S 181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                NON_TER
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 AAC DOT REPRESENTED TO THE SERVICE OF T
```

```
Job time : 36.7298 secs
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ပ
                                                                                                                                ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  TRTIASPKGPVIQMYTNVDKDLVGWPAPQGSRSLTPCTCGSSDLYLVTRHADVJPVRRRG 121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          DSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVCTRGVAKAVDFIPVESLETTMR 18;
                                                                                                                                                                                                                                                                                                              DSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVCTRGVAKAVDFIPVESLETTMR 181
                                                                                                                                                                                                                                                                                   62 TRTIASPKGPVIQMYTNVDKDLVGWPAPQGSRSLTPCTCGSSDLYLVTRHADV1PVRRKG 121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    2 APITAYAQQTRGLLGCIITSLTGRDKNOVEGEVQIVSTAAOTFLATCINGVCWTVYHGAG 61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    9
                                                                                                                                                                                                        5,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         2 APITAYAQOTRGLIGCIITSLIGRDKNOVEGEVOIVSTAAQIFLATCINGVCWTVYHGAG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Holland-Staley C.A., Kovari L.C., Golenberg E., Mayers D.L.;
"Genetic Diversity and response to IFN of the NS3 Protease Gene from Clinical Strains of the Hepatitis C Virus.";
Submitted (APR-201) to the EMBL/GenBank/DDBJ databases.
InterPro; IPR004109; HCV_NS3.
Pfam; PF02907; HCV_NS3; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Gaps
                                                                                                                                   Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                ..
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Length 181;
                                                                                Length 181,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1; Indels
                                                                                                                                Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      181 181
181 AA: 19116 MW; 9648867F49EBID43 CRC64;
                          181 AA; 19114 MW; 574AC47AE8AEE5D2 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  01-DEC-2001 (TrEMBLrel. 19, Created)
01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
                                                                          Score 936; DB 12;
Pred. No. 9.6e-88;
3; Mismatches 0;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              98.4%; Score 938; DB 12; 98.3%; Pred. No. 9.6e-68;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       181 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 2; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       PRT;
                                                                          98.48;
98.38;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Matches 178; Conservative
                                                                                                                                Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               NS3 protease (Fragment). Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  [1]
SEQUENCE FROM N.A.
                                                                                                     Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     NCBI_TaxID-11103;
181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                S 181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               5 182
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         S 182
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      S 181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                STRAIN-Pt.11
                                                                                                        Best_Local Sim
Matches 178;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Protease.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       62
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          122
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               182
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                   122
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         182
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Query Match
                          SEQUENCE
                                                                          Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              NON_TER
NON_TER
NON_TER
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Q91RT3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 RESULT 15
091RT3
10 01-DE
DT 0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ΟŻ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          a
                                                                                                                                                                                                                      q
                                                                                                                                                                                                                                                                                                                           q
                                                                                                                                                                                                                                                                                                                                                                                                                                 Ω
  FT
                                                                                                                                                                                   ö
                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
```

Search completed: August 30, 2003, 19:18:17

```
GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.
```

using sw model protein search, OM protein

August 30, 2003, 19:02:22 ; Search time 14.9789 Seconds (without alignments) 1168.452 Million cell updates/sec Run on:

US-09-965-594-1 953 Title: Perfect score: Sequence:

1 MAPITAYAQQTRGLLGCIII........GVAKAVDFTPVESLETIMRS 182

BLOSUM62 Gapop 10.0 , Gapext 0.5 Scoring table:

96168682 residues 283308 segs, Searched:

283308 Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0 Maximum DB seq length: 200000000

Post-processing: Minimum Match 0% Maximum Match 100% Listing first 45 summaries

Database :

PIR\_76:\*
1: pir1:\*
2: pir2:\*
3: pir3:\*
4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SANWYC3  Genome polypr   SANWYC4  Genome polypr   GNWYC4  Genome polypr   GNWYC4  Genome polypr   GNWYC4  Genome polypr   A45573  Genome polypr   A45573  Genome polypr   JO1303  Genome polypr   GNWYB  Genome polypr   GNWYB  Genome polypr   JO1303  Genome polypr   JO1304  Genome polypr   JO13	Result No.	Score			80	SUMMARIES	۰.
927 97.3 3011 1 GNWVCH 891 93.6 3010 1 GNWVTW 891 93.5 3010 1 GNWVTW 884 90.7 3010 1 GNWVCJ 884 90.7 3010 1 GNWVCJ 90.4 14.7 3010 1 SLB030 1 712 74.7 3013 1 JCS620 1 712 74.7 3013 1 JCS620 1 265.5 28.1 3005 2 TOB841 2 267.5 28.1 3005 2 TOB841 4 8 8 6 6 495 2 133184 8 8 6 495 2 173184 5 8 8 6 2 8 17384 7 8 8 2 8 1 398 2 B71284 7 8 8 2 8 1 398 2 B71284 7 7 8 8 1 398 2 B71284 7 7 8 8 2 2000 2 TOB175 7 7 8 8 6 8 2 END 2 TOB175 7 7 8 8 6 8 2 END 2 TOB175 7 7 8 8 6 8 2 END 2 TOB175 7 7 8 8 6 8 2 END 2 TOB175 7 7 8 8 8 8 8 8 8 2 END 2 TOB175 7 7 8 8 8 8 8 8 8 2 END 2 TOB175 7 7 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8	7 7	943 939		3011		GNWVC3 S40770	
4 892 93.6 3010 1 CNWVTW 5 887 93.6 3010 1 CNWVTW 7 884 92.8 3010 1 CNWVTC 884 92.8 3010 1 A45573 8 864 92.8 3010 1 A45573 8 864 92.8 3010 1 A45573 8 90.7 3010 1 S18030 0 714 74.9 3013 1 JC5620 0 714 74.9 3013 1 JC5620 1 712 267.5 281 2 3013 1 JC09831 2 255.5 26.8 2970 2 T08839 4 87 91 209 2 H393144 8 81.5 8.6 495 2 H39383 6 495 2 B71360 7 81.5 8.6 495 2 B71360 7 81.5 8.6 495 2 B71360 7 81.5 8.6 495 2 B71364 7 81.5 8.6 495 2 B71364 7 81.5 8.6 495 2 B71364 7 81.5 8.7 476 2 T88399 7 8.5 8.2 981 2 T18234 7 7 5 8.1 904 2 A84212 7 7 5 8.1 904 2 A84212 7 7 8 8.1 5 8.6 2 A8421 7 7 8 8.8 2 A8421 7 7 8 7 8 8 8 8 8 2 C81200	٣	927		3011	7	GNWVCH	
5 891 93.5 3010 1 GNWVTC 884 92.8 3010 1 AMVCJ 884 92.8 3010 1 AMVCJ 864 92.8 3010 1 AMVCJ 864 92.7 3010 1 S18030 9 804 80.7 3014 1 JC5620 714 74.9 3013 1 JC01303 2 257.5 28.8 2970 2 T08839 8 7 9.1 209 2 E193144 8 7 8.5 8.6 495 2 E139383 8 8 6 495 2 E139383 8 8 6 495 2 E139383 7 8 8 6 495 2 E139383 8 8 6 495 2 E139383 8 8 6 495 2 E139383 7 8 8 2 E1384 7 8 8 2 239 2 H89966 7 7 8 8 1 398 2 E17284 7 7 8 8 1 200 2 T06118 4 7 6 5 8 0 240 1 CPBOA3 7 7 8 8 1 240 2 T06118 7 7 8 8 1 240 1 CPBOA3 7 7 8 8 1 240 2 T06118 7 7 8 8 1 2 S70401 7 7 8 8 1 2 S70401	₹	892		3010	٦	GNWVTW	
6 887 93.1 3010 1 ANAVCJ 864 90.7 3010 1 S18030 9 864 90.7 3010 1 S18030 1 712 74.7 3013 1 JOH303 2 267.5 28.1 3005 2 T08841 2 267.5 28.1 3005 2 T08841 4 87 9.1 209 2 H8339 4 82 8.6 495 2 T08839 8 12 8.6 495 2 T18334 5 82 8.6 495 2 H8399 7 8.5 8.2 981 2 H83996 7 8.5 8.2 981 2 H83996 1 77.5 8.1 398 2 H8376 1 77.5 8.1 398 2 H8375 1 75.5 8.0 270 2 T06118 7 7.5 8.1 398 2 H8375 7 7.5 8.1 398 2 B71284 7 7.5 8.1 398 2 B71284 7 7.5 8.1 200 2 T06118 7 7.5 8.1 398 2 B71284 7 7.5 8.0 270 2 T06118 7 7.5 8.0 270 2 T06118 7 7.5 8.0 270 2 T06118 7 7.5 8.0 2 20 20 20 20 20 20 20 20 20 20 20 20	S.	891		3010	-	GNWV7C	
884 92.8 3010 1 A45573 884 92.8 3010 1 A45573 9 104 14.4 3014 1 JC5620 0 714 74.9 3013 1 JCNWVJB 712 267.5 28.1 3005 2 T08839 4 82 8.6 2970 2 T08839 4 82 8.6 495 2 139383 5 82 8.6 495 2 139383 6 82 8.6 495 2 139383 7 8.5 8.6 495 2 139384 7 8.5 8.6 495 2 139384 7 8.5 8.2 981 2 T18234 0 78 8.2 981 2 T18234 1 77.5 8.1 904 2 A84212 7 7 8.1 904 2 A8421 7 6.5 8.0 868 2 H81775 7 7 8.1 904 2 A8421 7 6.5 8.0 868 2 H81775 7 7 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8	91	887		3010	٦,	GNWVCJ	
9 804 80.7 3010 1 318030 1 714 74.9 3033 1 JCNWVJB 1 712 267.5 28.1 3033 1 JCNWVJB 2 267.5 26.8 2970 2 T08831 8 7 911 209 2 H393144 8 8 6 495 2 H39383 6 81.5 8.6 495 2 H39396 7 81.5 8.6 495 2 H39396 7 8.5 8.2 981 2 T1824 1 7.5 8.1 904 2 A8421 7 7 8 8.1 904 2 A8421 7 7 8 8.1 904 2 A8421 7 7 8 8.1 904 2 T06118 4 75 7 8 1 904 2 T06118 7 7.5 8 1 904 2 T06118 7 7 8 8 2 8 2 8 2 8 17284 7 7 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8	۰ ٥	884		3010	٦,	A45573	
1 267.5 28.1 3033 1 GNNVJS 267.5 28.1 3033 1 JQ1303 255.5 26.8 2970 2 T08839 4 87 9.1 209 2 Ii83144 5 82 8.6 495 2 139383 8 82 8.6 495 2 139383 8 82 8.6 495 2 139383 7 8.5 8.2 981 2 139376 7 8.5 8.2 981 2 18234 7 7 8 8.2 239 2 H89966 7 7 8 8.1 398 2 871284 7 7 8 8.1 290 2 T06118 4 75 8 9 240 1 CPBOA3 7 6.5 8 0 240 1 CPBOA3 7 7 8 1 8 15 2 570401 7 7 8 8 8 8 8 8 8 8 8 8 8 7 8 8 7 8	<b>0</b> 07	80 g		3010	٦,	S18030	
1 267.5 28.1 3033 1 JOJ303 2 255.5 26.8 1 3005 2 TO8841 4 87 9.1 209 2 H83144 5 82 8.6 495 2 TO8839 7 8.5 82 8.6 495 2 H33183 7 8.5 82 8.6 495 2 H3398 7 8.5 82 8.2 871360 7 7.5 8.1 299 2 H89966 1 77.5 8.1 398 2 B71284 7 7.5 8.1 398 2 B71284 7 7.5 8.1 200 2 TO6118 4 7.5 8.0 270 2 TO6118 5 74.5 7.8 868 2 END75 6 74.5 7.8 868 2 CR0003 7 74.5 7.8 868 2 CR0003	10	714		3033	1	GUNTUR	
2 267.5 28.1 3005 2 T08841 3 255.5 26.8 2970 2 T08839 5 82 8.6 455 2 H39383 6 82 8.6 495 2 H3360 8 1.5 8.6 495 2 H3360 7 8.5 8.2 620 2 H8399 7 8.5 8.2 981 2 T18234 0 78 8.2 981 2 T18234 1 77.5 8.1 2 98 2 B71284 2 77 8.1 904 2 A84212 3 76.5 8.0 270 2 T06118 4 76.5 8.0 270 2 H9175 5 75 7.9 240 1 CPBOA3 6 74.5 7.8 868 2 C81200 9 74.5 7.8 868 2 C81200 9 74.5 7.8 918 2 S70401 7 7.5 8 968 2 C81200	11	712		3033	7	J01303	
3 255.5 26.8 2970 2 T08839 4 87 9.1 209 2 H39144 6 82 8.6 455 2 H39184 7 81.5 8.6 495 2 H39189 7 78.5 8.2 981 2 T18234 0 78.5 8.2 981 2 T18234 1 77.5 8.1 904 2 M39965 1 77.5 8.1 904 2 M4175 3 76.5 8.0 270 2 H06118 4 75.7 8.1 904 2 M41775 7 7 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8	12	267.5		3005	7	T08841	polyprotein - dour
4 87 9.1 209 2 iiiiiiiiiiiiiiiiiiiiiiiiiiiiiiiii	13	255.5		2970	7	T08839	polyprotein - marm
5         82         8.6         452         2         139383           6         49         2         139383         8.6         495         2         131360           8         1.2         8.6         495         2         148399         8.2         148396         9.2         148396         9.2         148396         9.2         148396         9.2	14	87	9.1	209	?	1183144	probable aromatic
6 82 8.6 495 2 B71360 7 81.5 8.6 476 2 1848399 7 8.5 8.2 981 2 718234 0 78 8.2 981 2 718234 1 77 5 8.1 2 91 2 B71284 2 77 8.1 904 2 B71284 2 76.5 8.0 270 2 F06118 4 76.5 8.0 270 2 H05118 5 75 7.9 240 1 CPBOA3 6 74.5 7.8 868 2 C81200 7 74.5 7.8 868 2 C81200 9 74.5 7.8 968 2 C81200	15	82	8.6	452	~	139383	angio-associated m
7 81.5 8.6 476 2 148399 8 78.5 8.2 620 2 P83976 9 78.5 8.2 981 2 11824 1 7.5 8.1 398 2 B13284 1 7.5 8.1 398 2 B13284 2 77 8.1 904 2 B13284 3 76.5 8.0 270 2 T06118 4 6.5 8.0 270 2 T06118 5 75 7.9 240 1 CPBOA3 6 74.5 7.8 415 2 S70401 7 74.5 7.8 868 2 B1375 6 74.5 7.8 868 2 B1375 9 74.5 7.8 968 2 B130	16	82	8.6	495	~	B71360	hypothetical prote
8 78.5 8.2 620 2 189376 9 78.5 8.2 620 2 189376 1 77.5 8.1 398 2 1971284 1 77.5 8.1 398 2 1971284 2 76.5 8.0 270 2 706118 4 76.5 8.0 868 2 141775 4 76.5 8.0 868 2 141775 6 74.5 7.8 415 2 570401 7 74.5 7.8 868 2 6812 7 74.5 7.8 968 2 6812 8 74.5 7.8 968 2 68120 9 74.5 7.8 968 2 68120	17	81.5	9.	476	7	T48399	heat shock transcr
9 78.5 8.2 981 2 T18234 0 78 8.2 239 2 H39966 2 77 8.1 398 2 B71284 2 75.5 8.0 270 2 T06118 4 76.5 8.0 270 2 H06118 5 75.5 8.0 270 2 H06118 6 74.5 7.9 240 1 CPBOA3 6 74.5 7.8 415 2 S70401 74.5 7.8 868 2 C81200 9 74.5 7.8 911 2 JN0821	18	78.5	8.5	620	C)	F83976	cytochrome-c oxida
0 78 8.2 239 2 H89966 1 775 8.1 398 2 B71284 2 77 8.1 904 2 A84212 3 76.5 8.0 270 2 T06118 4 76.5 8.0 868 2 H81775 5 75 7.9 240 1 CPBOA3 6 74.5 7.8 415 2 S70401 7 74.5 7.8 868 2 C81200 9 74.5 7.8 911 2 JN0821	19	78.5	8.5	981	7	T18234	beta transducin ho
1 77.5 8.1 398 2 b71284 3 76.5 8.0 270 2 706118 4 76.5 8.0 240 1 CPBOA3 6 74.5 7.8 415 2 570401 7 74.5 7.8 868 2 H8175 6 74.5 7.8 8755 2 532441 7 74.5 7.8 868 2 CB1200 9 74.5 7.8 968 2 CB1200	20	78		239	~	н89966	serine proteinase
2 76.5 8.1 904 2 A84212 3 76.5 8.0 270 2 106118 4 76.5 8.0 868 2 H81775 5 7.9 240 1 CPBOA3 6 74.5 7.8 415 2 \$70401 7 74.5 7.8 868 2 C81200 9 74.5 7.8 911 2 JN0821	21	77.5	œ	398	7	871284	probable periplasm
3 76.5 8.0 270 2 706118 4 76.5 8.0 868 2 H81775 7 7.9 240 1 CPBOA3 6 74.5 7.8 415 2 870401 7 44.5 7.8 868 2 CB1200 9 74.5 7.8 911 2 JN0821	22	77	•	904	7	A84212	hypothetical prote
4 76.5 8.0 868 2 H81775 aconitate hyd 5 75 7.9 240 1 CPBOA3 procarboxypep 6 74.5 7.8 415 2 S23441 hypothetical 8 74.5 7.8 868 2 C81200 aconitate hyd 5 7.8 911 2 JN0821 transferrin-b	23	76.5		270	7	T06118	hypothetical prote
5 75 7.9 240 1 CPBOA3 6 74.5 7.8 415 2 S20401 7 74.5 7.8 755 2 S23441 8 74.5 7.8 868 2 CB1200 9 74.5 7.8 911 2 JN0821	74	76.5	•	868	~	H81775	aconitate hydratas
6 74.5 7.8 415 2 \$70401 7 4.5 7.8 755 2 \$23441 8 74.5 7.8 868 2 C81200 9 74.5 7.8 911 2 JN0821	25			240	-	CPBOA3	procarboxypeptidas
7 74.5 7.8 755 2 523441 8 74.5 7.8 868 2 C81200 9 74.5 7.8 911 2 JN0821	56	74.5		415	~	S70401	zona pellucida gly
8 74.5 7.8 868 2 CB1200 9 74.5 7.8 911 2 JN0821	27	74.5		755	7	S23441	hypothetical prote
9 74.5 7.8 911 2 JN0821	78	74.5	•	868	7	CB1200	aconitate hydratas
	53	74.5		911	~	JN0821	transferrin-bindin

hypothetical prote	protein F22G5.28	hypothetical prote	hypothetical prote	C4-dicarboxylate t	phosphate ABC tran	hypothetical prote	transferrin-bindin	transferrin-bindin	latent transformin	serine proteinase	hypothetical prote	htra-like serine p	chitinase (EC 3.2.	granzyme M (EC 3.4
C72705	G86208	F90978	S76033	HB2044	H84203	T04533	C81832	F81196	A55494	A89967	T49806	H97199	S65762	A55634
7 0	7	7	7	~	~	~	7	7	7	7	~	7	~	7
140	808	356	447	451	266	846	910	915	1820	239	354	433	535	249
7.8	7.8	7.7	7.7	7.7	7.7	7.7	7.7	7.7	7.7	7.7	7.7	7.7	7.7	7.6
74	74	73.5	73.5	73.5	73.5	73.5	73.5	73.5	73.5	73	73	73	73	72.5
30	32	33	34	35	36	37	38	5.5	0.5	41	42	43	44	45

## ALIGNMENTS

```
A; Motecule type: genomic RNA
A; Residues: 1577-1633 CCHA>
A; Residues: 1577-1634 CCHA>
A; Experimental source: 1solates E-b16
A; Experimental source: 1solates E-b16
A; Accession: p00404
A; Status: preliminary
A; Molecule type: genomic RNA
A; Residues: 1577-1633 CCH2>
A; Reperimental source: 1solates E-b17
C; Superfamily: hepatitis C virus genome polyprotein
C; Superfamily: hepatitis C virus genome p
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              A.Accession: A39166
A.Molecule type: mRNA
A.Residues: 1-3011 ccDo.
A.Residues: 1-3011 ccDo.
A.Residues: 1-3011 ccDo.
A.Cross-references: GB:M62321; NID:g329873; PIDN:AAA45676.1; PID:g329874
A.Cross-references: GB:M62321; NID:g329873; PIDN:AAA45676.1; PID:g329874
B.Chan, S.W.; McOmish, F.; Holmes, E.C.; Dow, B.; Peutherer, J.F.; Follett, E.; Yap
J. Gen. Virol. 73, 1131-1141, 1992
A.Title: Analysis of a new hepatitis C virus type and its phylogenetic relationship
A.Reference number: PQ0403
A.Recession: PQ0403
genome polyprotein - hepatitis C virus (strain HCV-1)
N;Contains: capsid protein C; envelope protein M; hepacivirin (EC 3.4.21.98) (nonst protein NS4a; nonstructural protein NS4b; nonstructural protein NS5
C;Species: hepatitis C virus
C;Accession: A39166; pO0403; pO0404
R;Choo, Q.L.; Richman, K.H.; Han, J.H.; Berger, K.; Lee, C.; Dong, C.; Gallegos, C.
Proc. Natl. Acad. Sci. U.S.A. 88, 2451-2455, 1991
A;Fille: Genetic organization and diversity of the hepatitis C virus.
A;Reference number: A39166; MUID:91172826; PMID:1848704
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 F.1616-1862/Product: nonstructural protein NS4a #status predicted <N4A>
F.15163-2013/Product: nonstructural protein NS4b #status predicted <N4B>
F.2014-3011/Product: nonstructural protein NS5 #status predicted <NS5>
F.196,209,234,305,325,417,423,430,448,476,532,540,556,576,623,645,1213,1255,2041,20
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Ouery Match 99.0%; Score 943; DB 1; Length 3011; Best Local Similarity 98.4%; Pred. No. 5.7e-81; Matches 179; Conservative 3; Mismatches 0; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          F;1316-1319/Region: DEXH motif
```

ò

us-09-965-594-1.rpr

COM

```
A.Note: meither amino acid nor nucleotide sequence is given
C.Superfamily: hepatitis C virus genome polyprotein
C.Superfamily: hyproduct: capsid protein M #Status predicted <CPC>
F:115.192.789/Product: major envelope protein E *status predicted <NSI>
F:390.729/Product: major envelope protein NSI *status predicted <NSI>
F:730.1006/Product: nonstructural protein NSI *status predicted <NSI>
F:730.1206/Product: nonstructural protein NSI *status predicted <NSI>
F:730.1307.1377/Region: nucleotide-binding motif B
F:7312-1317/Region: DEXH motif
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        F;1616-1862/Product: nonstructural protein NS4a #status predicted <N4A>
F;1863-2013-Product: nonstructural protein NS5# #status predicted <N4B>
F;2014-3011/Product: nonstructural protein NS5# status predicted <NS5>
F;2014-3011/Product: nonstructural protein NS5#status predicted <NS5>
F:196,209,234,305,325,417,423,430,448,476,532,540,556,576,623,645,1213,1255,2041,224
                                                                                                                                        genome polyprotein - hepatitis C virus (strain H)
N;Contains: capsid protein C; envelope protein M; hepacivirin (EC 3.4.21.98) (nonstr
protein NS4a; nonstructural protein NS4b; nonstructural protein NS5
                                                                                                                                                                                                                                                                                                                                                                                                                                                       A; Description: Genomic structure of the human prototype strain H of hepatitis C viru
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       A; Molecule type: genomic RNA
A; Residues: 1-3011 < INC>
A; Residues: 1-3011 < INC>
A; Cross-references: GB:M67463; NID:g329737; PIDN:AAA45534.1; PID:g329738
A; Cross-references: GB:M67463; NID:g329737; PIDN:AAA45534.1; PID:g329738
B; Inchauspe, G; Zebedee, S.; Lee, D.H.; Sugitani, M.; Nasoff, M.; Prince, A.M. Proc. Natl. Acad. Sci. U.S.A. 88, 10292-10296, 1991
A; Title: Genomic Structure of the human prototype strain H of hepatitis C virus: A; Contents: annotation
                                                                                                                                                                                                                                                                          A; Note: host Homo sapiens (man)
C; Date: 31-Dec-1992 #sequence_revision 31-Dec-1992 #text_change 19-Jan-2001
C; Accession: A36B14; A41546
R; Inchauspe, G:; Zebede, S.; Lee, D.H.; Sugitani, M.; Nasoff, M.; Prince, Psubmitted to GenBank, July 1992
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     97.3%;
96.2%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                175; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Species: hepatitis C virus
                                                                                                                                                                                                                                                   C; Species: hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            A; Reference number: A36814
A; Accession: A36814
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   RS 1207
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    RS 182
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       C:Accession: A40244
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Query Match
Best Local Si
Matches 175
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1206
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                đ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Sedicate polyprotein - hepatitis C virus
Michaelas: capsid protein C: envelope protein M: hepacivirin (EC 3.4.21.98) (nonstructu protein MS4s; nonstructural protein NS4b: nonstructural protein NS5becles: hepatitis C virus
C:Species: 19-May-2000 #sequence_revision 19-May-2000 #text_change i9-Jan-2001
C:Accession: S40770; PC1285
R:Gramoto, H.
A:Accession: S40770
A:Accession: PC1284; MUD:9221586; PIDN:BAA01582.1; PID:9221587
A:Accession: PC1284; MUD:91013116; PMID:2170712
A:Accession: PC1285
A:Accession:
                                                                                                                                        1145
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ن:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    LAPITAYAQQIRGLLGCIIISLIGRDKNQVEGEVQIVSTAAQTFLATCINGVCWTVYHGA 1085
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   120
1026 LAPITAYAQQTRGLLGCITTSLTGRDKNQVEGEVQIVSTAAQIFLATCINGVCWIVYHGA 1085
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   121 GDSRGSLLSPRPISYLKGSSGGPLICPAGHAVGIFRAAVCTRCVAKAVDFIPVESLETIM 186
                                                                                                                                                                                                              121 GDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVCTRGVAKAVDFIPVESLETTM 180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      09
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1 MAPITAYAQQTRGLLGCIITSLTGRDKNQVEGEVQIVSTAAQTFLATCINGVCWTVYHGA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           GTRTIASPKGPVIQMYTNVDKDLVGWPAPQGSRSLTPCTCGSSDLYLVTRHADVIPVRRR
                                                                                                           GTRIIASPKGPVIQMYTNVDKDLVGWPAPQGSRSLIPCTCGSSDLYLVTRHADVIPVRRR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                F;1616-1862/Product: nonstructural protein NS4a #status predicted <N4A>F;1863-2013/Product: nonstructural protein NS4b #status predicted <N4B>F;2014-3011/Product: nonstructural protein NS5 #status predicted <NS5>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ċ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Length 3011;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Score 939; DB 1;
Pred. No. 1.4e-80;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               3; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             98.5%;
97.8%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ||
| RS 1207
                                                                                                                                                                                                                                                                                                                                                                                                                      RS 1207
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   RS 182
                                                                                                                                                                                                                                                                                                                                                   181 RS 182
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Matches 178;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1026
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   181
                                                                                                                                                                                                                                                                                                                                                                                                                          1206
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1206
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Query Match
Best Local 9
                                                                                                                                 qq
                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                Š
                                                                                                                                                                                                                                                                                                                                                                                                                      a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ŝ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ô
                                                                                                                                                                                                           ö
                                                                      õ
```

```
genome polyprotein - hepatitis C virus (strain faiwan)
N;Contains: capsid protein C; envelope protein M; hepacivirin (EC 3.4.21.98) (nonstriprotein NS4a; nonstructural protein NS4b; nonstructural protein NS5
                                                                                                                                                                  1085
                                                                                                                                                                                                                                                                                                                                                              ö
                                                                                                                                                                                                                                                                                                                                 121 GDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVCTRGVAKAVDFIPVESLETTM 180
                                                                                                                                                                                                                      61 GTRIIASPKGPVIQMYINVDKDLVGWPAPQGSRSLIPCTCGSSDLYLVTRHADVIPVRRR 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     A:Note: host Homo sapiens (man)
C:Date: 31-Dec-1992 *sequence_revision 31-Dec-1992 *text_change 19-Jan-2001
                                                                                                                                           1 MAPITAYAQQTRGLLGCIITSLTGRDKNQVEGEVQIVSTAAQTFLATCINGVCWTVYHGA
                                                            Gaps
                                                          .
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   R;Chen, P.J.; Lin, M.H.; Tai, K.F.; Liu, P.C.; Lin, C.J.; Chen, D.S. virology 188, 102-113, 1992
  Length 3011;
                                                          Indels
Score 927; DB 1;
Pred. No. 1.9e-79;
                                                          4; Mismatches
```

```
F;196,209,234,250,305,325,417,423,430,448,532,540,556,576,623,645,1213,1255,2041,20
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             61 GIRIIASPKGPVIQMYINVDKDLVGWPAPQGSRSLIPCTCGSSDLYLVTRHADVIPVRRR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 17; Mismatches
                                                                                                                                                      Mismatches
                                                                                                                                                  15;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            93.1%;
87.9%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Conservative
                                                                                                                                                      Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Similarity
                                                                                                               Local Similarity
es 162; Conserves
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        RS 1207
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       181 RS 182
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     160;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1026
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1206
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        121
                                                                                Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Local
                                                                                                                                                      Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                             a
                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                       a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ô
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     genome portprotein . Hepatitis C. VIIUS

N; Contains: capsid protein C. envelope protein M; hepacivirin (EC 3.4.21.98) (nonstructu protein NS42; nonstructural protein NS5

C; Accession: A3846; nonstructural protein NS4b; nonstructural protein NS5

C; Accession: A3846; Mile Silve, I.; Manabe, S.; Murakami, S.; Fujita, J.; Cnishi, F.; J. Virol. 65, 1105-1113, 1991

A; Title: Structure and organization of the hepatitis C virus genome isolated from human A; Reference number: A38465; MID:91140698; PMID:1847440

A; Reference number: A38465; MID:91140698; PMID:1847440

A; Reference number: A38465; MID:91140698; PMID:1847440

A; Residues: 1-3010 crax

C; Superfamily: hepatitis C virus genome polyprotein: qivcoprotein: hydrolase; nonstructural protein NS; fistatus predicted cvas

F; 130-1317, Region: nucleotide-binding motif B

F; 1316-1319, Region: nucleotide-binding motif B

F; 1316-1319, Region: nucleotide-binding motif B

F; 1316-1319, Region: nucleotide-binding motif B

F; 1316-1300, Region: nucleotide-bi
                                                                                                                                                             A: Cross-references: CB:MB4754
C: Superfamily: hepatitis C virus genome polyprotein
C: Superfamily: hepatitis C virus genome polyprotein
C: Reywords: ATP: capsid protein: envelope protein: quycoprotein; hydrolase; nonstructural
F: 1-115/Product: capsid protein C *status predicted <CPC>
F: 116-191/Product: capsid protein M *status predicted <CPC>
F: 116-191/Product: major envelope protein B *status predicted <NES>
F: 130-729/Product: nonstructural protein NS1 *status predicted <NS1>
F: 130-729/Product: nonstructural protein NS1 *status predicted <NS2>
F: 1200-1515/Product: hepacivirin *status predicted <NS3>
F: 1316-1317/Region: nucleotide-binding motif B (P-loop)
F: 1316-1317/Region: DEXH motif
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            F:1616-1862/Product: nonstructural protein NS4a #status predicted <N4A>
F:1863-2013/Product: nonstructural protein NS4b #status predicted <N4B>
F:2013/Product: nonstructural protein NS5 #status predicted <NS5>
F:2014-3010/Product: nonstructural protein NS5 #status predicted <NS5>
F:196,209,233,234,250,305,325,417,423,430,448,532,540,556,576,623,645,1213,1255,2041,207
A;Title: The Taiwanese hepatitis C virus genome: sequence determination and mapping the A;Reference number: A40244; MUID:92230206; PMID:1314449
A;Accession: A40244
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1086 GSKTLAGPKGPITQMYTNVDQDLVGWHAPQGARSLTPCTCGSSDLYLVTRHADVIPVRRR 1145
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              121 GDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVCTRGVAKAVDFIPVESLETTM 180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ၀
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1 MAPITAYAQQTRGLLGCIITSLTGRDKNQVEGEVQIVSTAAQTFLATCINGVCWTVYHGA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     93.6%; Score 892; DB 1; Length 3010; 90.1%; Pred. No. 4.1e-76;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              12; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          genome polyprotein - hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Matches 164; Conservative
                                                                                                               A; Molecule type: genomic RNA A; Residues: 1-3010 <CHE>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             RS 1207
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           RS 182
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
```

procein Nasa, nonstructural protein NS4D; nonstructural protein NS5 (Species: Nepatitis C virus
C; Species: Nepatitis C virus
C; Date: 30 - vul 1922 sequence\_revision 30 - Jun-1992 \*text\_change 19 - Jan-2001
C; Accession: A3923; ps0086
B; Kato, N.; Hijikata, M.; Octsuyama, Y.; Nakagava, M.; Ohkoshi, S.; Sugimura, T.; S
Proc. Natl. Acad. Sci. Us.A. 87, 872
A; Title: Molecular cloning of the human hepatitis C virus genome from Japanese patl
A; Receiver contumber: A3923; MUD:91088550; PMID:2175903
A; Molecula type: genomic RNA
A; Residues: 1-3010 «KAT>
A; Caross-reference: CB: 1990208; NID:9221610; PIDN:BAA14233.1; PID:9221611
B; Kato, N.; Ohkoshi, S.; Shimotohno, K.
A; Testerence number: P80085
A; Title: Japanese isolates of the non-A, non-B hepatitis viral genome show sequence
A; Reference number: P80086
A; Molecule type: genomic RNA
A; Residues: 2650-2707 «KA2>
A; Residues: genome polyprotein - hepatitis C virus (strain J) N;Contains: capsid protein C; envelope protein M; major envelope protein E; nonstru protein NS4a; nonstructural protein NS4b; nonstructural protein NS5 1145 ö 1 MAPITAYAOOTRGLLGCIITSLTGRDKNOVEGEVOIVSTAAOTFLATCINGVCWTVYHGA 1086 GSKTLAAPKGPITQMYTNVDQDLVGWPKPPGARSLTPCTCGSSDLYLVTRHADVIPVRRR GDSRGSLLSPRPISYLRGSSGGPLLCPAGHAVGIFRAAVCTRGVAKAVDFIPVESLETTM 61 GTRIIASPKGPVIQMYTNVDKDLVGWPAPQGSRSLTPCTCGSSDLYLVTRHADVIPVRRR 1 MAPITAYAQQTRGLLGCIITSLTGRDKNQVEGEVQIVSTAAQTFLATCINGVCWTVYHGA Length 3010; 5; Indels Score 887; DB 1; Pred. No. 1.2e-75; 93.5%; Score 891; DB 1; 89.0%; Pred. No. 5.1e-76;

ŭ,

single

from a

ŭ

```
A; Experimental Source: isolate JI from an individual R; Honda, M.; Kaneko, S.; Unoura, M.; Kobayani, K.; Murakami, S. Arch. Virol. 128, 163-169, 1993
A; Title: Sequence analysis of putative structural regions of hepatitis C virus isolai A; Réference number: A48332; MUID:93119270; PMID:8380322
A; Récession: $33570
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         F;1616-1862/Product: nonstructural protein NS4a *status predicted <N4A>
F;1863-32013/Product: nonstructural protein NS4b *status predicted <NNB>
F:2014-3010/Product: nonstructural protein NS5 *status predicted <NS5-
F:196,209,234,250,305,417,423,448,532,540,556,576,623,645/Binding site: carbohydrate
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          A;Molecule type: genomic RNA
A;Residues: 1-547,'T',549-621,'V',623-624,'S',626-652,'DL',655-761,'T',763-782 <HOW>
A;Cross-references: EMBL:X61591
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  A.Note: this sequence is inconsistent with the nucleotide translation
A.Note: the authors translated the codon AGG for residue 43 as Pro, TGG for residue
as Trp, and TrC for residue 771 as Ser
A.Note: sequence extracted from WiBI Dackbone (NCBIN:121747, NCBIP:121748)
C.Superfamily: hepatitis C virus genome polyprotein
C.Superfamily: hepatitis C virus
C.Su
                      C:Date: 19-May-2000 #sequence_revision 19-May-2000 #text_change 23-Mar-2001 C:Date: 19-May-2000 #sequence_revision 189059
Accession: $180030; $33570; A48329
R:Honda, M.; Kaneko, S.; Masashi, U.; Kobayashi, K.; Murakami, S. submitted to the EMBL Data Library, September 1991
A:Boscription: A whole genome of hepatitis C virus cDNA was isolated from a A;Reference number: $18028
A;Accession: $18030
                                                                                                                                                                                                                                                                      A; Molecule type: genomic RNA
A; Residues: 1-3010 <HON>
A; Cross-references: EMBB:X61596; NID:959478; PIDN:CAA43793.1; PID:959479
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             F;1316-1319/Region: DEXH motif F;1616-1862/Product
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Best_Local Similarity
Matches 158; Conserv
A; Variety: isolate JK1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              RS 1207
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               RS 182
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1206
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             RESULT 9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  3C5620
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   δy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 N:Contains: caps1d protein C; envelope protein M; hepacivirin (EC 3.4.21.98) (nonstructu protein NS4a: nonstructural protein NS4b; nonstructural protein NS5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                              genome polyprotein - hepatitis C virus (strain JT) N:Contains: capsid protein C; envelope protein M: hepacivirin (EC 3.4.21.98) (nonstructu protein NS4a; nonstructural protein NS4b; nonstructural protein NS5 C; Species: hepatitis C virus C; Species: hepatitis C virus C; Date: 19-May-2000 #sequence_revision 19-May-2000 *text_change 19-Jan-2001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      R;Tanaka, T.; Kato, N.; Nakagawa, M.; Ootsuyama, Y.; Cho, M.J.; Nakazawa, T.; Hijikata, Virus Res. 23 39-53, 1992
Virus Res. 23 39-53, 1992
A;Title: Molecular cloning of hepatitis C virus genome from a single Japanese carrier: 8.
Reference number: A45573 MUID: 92295714; PMID: 1318627
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1086 GSKTLAGPKGPITQMYTNVDQDLVGWHAPPGARSLTPCTCGSSDLYI/VTRHADVIPVRRR :1145
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1146 GDGRGSLLSPRPVSYLKGSSGGPLLCPSGHAVGIFRAAVCTRGVAKAVDFIPVESMETIM 1205
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ;
0
                                                                                                                                        GDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVCTRGVAKAVDFIPVESLETTM 180
                                                                                                       121 GDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVCTRGVAKAVDFIPVESLETTM 180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           9
                             1086 GSKTLAGPKGPITQMYTNVDQDLVGWPAPPGARSMTPCTCGSSDLYLVTRHADVVPVRRR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1 MAPITAYAQQTRGLLGCIITSLTGRDKNQVEGEVQIVSTAAQTFLATCINGVCWTVYHGA
Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         <N4A>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              F.1616-1862/Product: nonstructural protein NS4a *status predicted <N4A: F;1863-2013/Product: nonstructural protein NS4b *status predicted <N4B: F;2014-3010/Product: nonstructural protein NS5 *status predicted <NS5>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Length 3010;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         92.8%; Score 884; DB 1; LC 92.8%; Pred. No. 2.38-75; 6.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   polyprotein - hepatitis C virus (isolate JK1)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         A;Molecule type: DNA
A;Residues: 1-3010 <TAN>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Query Match
Best Local Similarity
Matches 161; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           A;Status: preliminary
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 RS 1207
                                                                                                                                                                                                                                                                                                           RS 1207
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               RS 182
                                                                                                                                                                                                                                            RS 182
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    C; Accession: A45573
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               181
                                                                                                                                                                                                                                            181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                æ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       genome
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              S18030
                                                                                                                                                                                                                                                                                                                                                                                                                       RESULT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                      셤
                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                    g
                                                                                                       ö
```

```
genome polyprotein - hepatitis C virus (isolate EUR1480)
N:Contains: capsid protein C; envelope protein M; hepacivirin (EC 3.4.21.98) (nonstruprotein Nsds; nonstructural protein NS5
C;Species: hepatitis C virus
C;Species: hepatitis C virus
C;Date: 19-May-2000 *sequence_revision 19-May-2000 *text_change 19-Jan-2001
C;Accession: JC5620
R;ChamberLain, R.W.; Adams, N.J.; Taylor, L.A.; Simmonds, P.; Elliott, R.M.
Biochem. Biophys. Res. Commun. 236, 44-49, 1997
A;Title: The complete coding sequence of hepatitis C virus genotype 5a, the predomins A;Reference number: JC5620; WUID:97366593; PMID:9223423
                                                                                                                                                                                                                 1085
                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ö
                                                                                                                                                                                                                                                                                        120
                                                                                                                                                                                                                                                                                                                                                                                                                               121 GDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVCTRGVAKAVDFIPVESLETTM 180
                                                                                                                                                                                   1086 GSKILAGPKGPINOMYTNVDQDLVGWQAPSGAASLTPCTYGSSDLYLVTRHADVIPVRRR
                                                                                                                                                 1 MAPITAYAQQTRGLLGCIITSLTGRDKNQVEGEVQIVSTAAQTFLATCINGVCWTVYHGA
                                                                                                                                                                                                                                                                                        GTRIIASPKGPVIQMYINVDKDLVGWPAPQGSRSLIPCTCGSSDLYLVTRHADVIPVRRR
                                                                              Gaps
   Length 3010;
                                                                       Indels
90.7%; Score 864; DB 1; E. 86.8%; Pred. No. 1.9e-73; iive 15; Mismatches 9;
```

C; Species: hepatitis C virus

```
R; Okamoto, H; Okada, S.; Sugiyama, Y.; Kurai, K.; Ilzuka, H.; Machida, A.; Miyakaw, J. Gen. Virol. 72, 2697-2704, 1991
A; Tille: Nucleotide sequence of the genomic RNA of hepatitis C virus isolated from A; Reference number: J01303; MUID:92044440; PMID:1658196
A; Accession: J01303
A; Molecule type: genomic RNA
A; Molecule type: genomic RNA
A; Residues: 1-3033 < OKA>
A; Cross-references: GB:D00944; NID:9221650; PIDN:BAA00792.1; PID:9221651
A; Residues: 1-3033 < OKA>
A; Cross-references: GB:D00944; NID:9221650; PIDN:BAA00792.1; PID:9221651
A; Residues: 1-3033 < OKA>
A; Cross-references: GB:D00944; NID:9221650; PIDN:BAA00792.1; PID:9221651
C; Superfamily: hepatitis C virus genome polyprotein
C; Superfamily: hepatitis D virus predicted < NED>
F; J16-191/Product: enpostructural protein NS1 *status predicted < NS2>
F; J131-1619/Product: nonstructural protein NS4 *status predicted < NAD>
F; J186-2017/Product: nonstructural protein NS5 *status predicted < NOS>
F; J186-2017/Product: nonstructural protein NS5 *status predicted < NOS>
F; J186-209, 234, 305, 325, 417, 423, 430, 448, 477, 534, 542, 558, 578, 627, 649, 1091, 1217, 1259, 20
                                                       genome polyprotein - hepatitis C virus (isolate HC-J6)
N'Contains: capsid protein C; envelope protein M: hepacivirin (EC 3.4.21.98) (nonst protein NS4s; nonstructural protein NS4b; nonstructural protein NS5 C;Species: hepatitis C virus
C;Date: 19-May.2000 *sequence_revision 19-May-2000 *text_change 17-Nov-2000 C;Accession: JQ1303
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 61 GTRTIASPKGPVIQMYTNVDKDLVGWPAPQGSRSLTPCTCGSSDLYLVTRHADVIPVRRR 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         121 GDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVCTRGVAKAVDFIPVESLETTM 180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1 MAPITAYAQQTRGLLGCIITSLTGRDKNQVEGEVQIVSTAAQIFLATCINGVCWTVYHGA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ;
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Length 3033;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   27; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        74.9%; Score 714; DB 1; 70.9%; Pred. No. 3.5e-59;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           26; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Best Local Similarity 70.9%;
Matches 129; Conservative 2
A; Molecule type: mRNA
A: Residues: 2678-2729 <KAT>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1210 RT 1211
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       RS 182
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       RESULT 11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               JQ1303
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                              A; Cross-references: GB:Y13184
A; Experimental source: genotype 5a, which predominates in South Africa
A; Retimental source: genotype 5a, which predominates in South Africa
A; Note: the translation of the nucleotide sequence is not complete in this paper
C; Superfamily: hepatitis C virus genome polyprotein
C; Keywords: ATP: glycoprotein; hydrolase; nucleotide binding; P-loop; polyprotein: C; Keywords: ATP: glycoprotein; hydrolase; nucleotide CEPM>
F; 116-191/Product: capsid protein C *status predicted <EPM>
F; 121-1389/Product: major envelope protein Ms! *status predicted <MED>
F; 384-408/Region: hypervariable *status predicted <NSI>F; 330-730/Product: nonstructural protein NSI *status predicted <NSI>F; 1231-1238/Region: nucleotide-binding motif A (P-loop)
F; 1317-1320/Region: nucleotide-binding motif A (P-loop)
F; 1317-1320/Region: DEXH motif C P-loop)
F; 1317-1320/Region: nucleotide-binding motif B
F; 1317-134/Product: nonstructural protein NS4 *status predicted <N4B>
F; 2015-3014/Product: nonstructural protein NS5 *status predicted <NAB>
F; 2015-3014/Product: nonstructural protein NS5 *status predicted <NS5>
F; 2015-3014/Product: nonstructural protein NS6 *status predicted <NS5>
F; 2015-3014/Product: nonstructural protein NS6 *status predicted <NS6>
F; 2015-3014/Product: nonstructural protein NS6 *status predicted <NS6</p>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           A; Experimental source: isolate E-bl2
R; Kato, N.; Octsuyama, Y.; Ohkoshi, S.; Nakazawa, T.; Mori, S.; Hijikata, M.; Shimotohnd
Blochem. Biophys. Res. Commun. 181, 279-285, 1991
A; Title: Distribution of plural HCV types in Japan.
A; Reference number: PQ0554; MUID:92068204; PMID:1720309
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       C;Species: hepatitis C vitus
C;Species: hepatitis C vitus
C;Species: hepatitis C vitus
C;Date: 31-Dec-1992 *sequence_revision 31-Dec-1992 *text_change 19-Jan-2001
C;Accession: A40250; p00397; p00559
R;Okamoto, H.: Kurai, K.; Okada, S.I.; Yamamoto, K.; Lizuka, H.; Tanaka, T.: Fukuda, S.; Vitology 188, 331-341, 1992
A;Title: Full-length sequence of a hepatitis C virus gcnome having poor homology to repd A;Reference number: A40250; MUID:92230232; PMID:1314459
A;Accession: A40250
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     genome polyprotein - hepatitis C virus (strain iiC-38)
N;Contains: capsid protein C; envelope protein M; hepacivirin (EC 3.4.21.98) (nonstructu
protein NS4a; nonstructural protein NS4b; nonstructural protein NS5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          A; Molecule type: genomic RNA
A; Residues: 1-3033 < COKA>
A; Cross-references: GB:DD0989; GB:D01221; NID:g221608; PIDN:BAA01761.1; PID:g221609
A; Cross-references: GB:DD088; GB:D0221; NID:g221608; PIDN:BAA01761.1; PID:g221609
B; Chan, S.W.: McOmish, F.: Holmes, E.C.; Dow, B.: Peutherer, J.F.; Follett, E.; Yap, P.I.
J. Gen. Virol. 73, 1131-1141, 1992
A; Title: Analysis of a new hepatitis C virus type and its phylogenetic relationship to e
A; Reference number: PQ0393; MUID:92268871; PMID:1316939
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      GTRTIASPKGPVIQMYTNVDKDLVGWPAPQGSRSLTPCTCGSSDLYLVTKHADVIPVRRR 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    09
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1 MAPITAYAQQTRGLLGCIITSLTGRDKNQVEGEVQIVSTAAQTFLATCINGVCWTVYHGA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            F;2210-2249/Region: interferon sensitivity determining *status predicted
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ;
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Length 3014;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         16; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              84.4%: Score 804; DB 1; 79.1%; Pred. No. 9.5e-68;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     22; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             A;Molecule type: genomic RNA
A;Residues: 2678-2754 <CHA>
A;Cross-references: DDBJ:D10134
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Matches 144; Conservative
                               A;Residues: 1-3014 <CHA>
A;Cross-references: GB:Y13184
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Local Similarity
   Molecule type: mRNA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                RS 1208
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     RS 182
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1207
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             අු
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        a
```

9

33;

Score 255.5; DB 2; Pred. No. 1.4e-15; 6; Mismatches 68;

```
A;Cross-references: EMBL:AF023424; NID:92828597; PIDN:AAC40501.1; PID:92828598 C;Superfamily: hepatitis C virus genome polyprotein C;Keywords: polyprotein
                                                                                                                                                           Query Match 26.8%; Score 255.5; Di
Best Local Similarity 30.1%; Pred. No. 1.4e-
Matches 59; Conservative 36; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               C; Species: Homo sapiens (man)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    67 SPKGP-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    106
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 109
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     122
                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        d
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 d
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              õ
                                                                                                                                                                                                                                                                                                                                                                                                                            õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Deprotein - douroucouli hepatitis GB virus A Cipatotein - douroucouli hepatitis GB virus A Cipatotein - douroucouli hepatitis GB virus A Cipatoteis: douroucouli hepatitis GB virus A Cipatotei Cipatotei
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       R.Erker, J.C.: Desai, S.M.: Leary, T.P.: Chalmers, M.L.: Montes, C.C.; Mushahwar, I.K. J. Gen. Virol. 79, 41-45, 1998
A.Title: Genomic analysis of two GB virus A variants isolated from captive monkeys. A.Reference number: 216486; MUID:98120818; PMID:9460920
A.Recession: T08839
                                                                                                                                                                                                                                                                                               .,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1038 ARTLAGPVGPVNCRWWSPSDDVAVYPLPSGASCLEPCKCGTQSVWCIRN--DGALCHGRL 1095
                                                                      <u>.</u>
                                                                                                                                                                                                                                                                                                                                                                                          GDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVCTRGVAKAVDFIPVESLETTM 180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               62 TRIIASPKGPVIQMYTNVDKDLVGWPAPQGSRSLTPCTCGSSDLYLVTRHADVIPVRRRG 121
                                                                                                                                   09
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      2 APITAYAQQTRGLLGCIITSLTGRDKNQVEGEVQIVSTAAQTFLATCINGVCWTVYHGAG 6:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       polyprotein - marmoset hepatitis GB virus A
C;Species: marmoset hepatitis GB virus A
C;Date: 23-Jul-1999 #sequence_revision 23-Jul-1999 #text_change 17-Nov-2000
C;Accession: 100839
                                                                                                                                   1 MAPITAYAQQTRGLLGCITTSLTGRDKNQVEGEVQIVSTAAQTFILAICINGVCWTVYHGA
                                                                                                                                                                                                                                                                GTRIIASPKGPVIQMYTNVDKDLVGWPAPQGSRSLIPCTCGSSDLYLVTRHADVIPVRRR
                                                                          Gaps
                                                                   0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Length 3005;
   Length 3033;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Indels
                                                                   Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  28.1%; Score 267.5; DR 2; 33.1%; Pred. No. le-16;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     78;
Ouery Match 74.7%; Score 712; DB 1; L
Best Local Similarity 69.8%; Pred. No. 5.4e-59;
Matches 127; Conservative 29; Mismatches 26;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     32; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             A;Status: translated from GB/EMBL/DDBJ
A;Molecule type: genomic RNA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Query Match 28.1%
Best Local Similarity 33.1%
Matches 60; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                RS 1211
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      S 1155
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       181 RS 182
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      S 175
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      175
                                                                                                                                                                                                                                                                                                                   q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ά
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               qq
                                                                                                                                                                                             q
                                                                                                                                                                                                                                                                ŏ
                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                a
```

```
probable aromatic acid decarboxylase PA4019 [imported] - Pseudomonas aeruginosa (str. C.Species: Pseudomonas aeruginosa (str. C.Scover, C.K.; Plaa, X.O.; Erwin, A.L.; Mizoguchi, S.D.; Warrener, P.; Hickey, M.J. adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; I. J. Lory, S.; Olson, W.V. Nature 406, 959-964, 2000
A;Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic p. A;Reference number: A82950; MUID:20437337; PMID:10984043
A;Accession: H83144
A;Status: preliminary
A;A;Medecule type: DNA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     A; Residues: 1-09 <STO>
A; Cross-references: GB:AE004818; GB:AE004091; NID:g9950200; PIDN:AAG07406.1; GSPDB:GI
A; Experimental source: strain PA01
C; Genetics:
C; Genetics:
C; Superfamily: dedF protein
                                                                            970 APVIVIH-HHGKGFFGVVKTSMTGWDETEHVGNVVVLGTSTTRSMGTCVNGVMYTTYHGSN 1028
                                                                                                                                                                                                                                                                                                                                                            ---VIQMYTNVDKDLVGWPAPQGSRSLTP----CTCGSSDL 105
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     54 LPAKPQAMQAFLTEYCGAAAGQIRVFGQND-----WMAPPASGSSAPNAMVICPCSTGTL 108
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          -----YLVTRHADVIPVRRRGDSRGSLLSPR--PIS-----YLKGSSGGPLLCPA 148
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SAVATGACNNLIERAADVALKER----RPLVLVPREAPFSSIHLENMLKLSNLGAVILPA 164
                                                                                                                                                             62 TRIIASPKGPVIQMYTNVDKDLVGWPAPQGSRSLTPCTCGSSDLYLVTRHADVIPVRRG 121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           8 AQQIRGLLGCIITSLIGRDKNQVEGEVQ-IVSTAAQIFLATCINGVCWTVYHGAGIRTIA 66
2 APITAYAQOTRGLLGCIITSLTGRDKNOVEGEVQIVSTAAQTFLATCINGVCWTVYHGAG 61
                                                                                                                                                                                                       DSRGSLLS-----PRPISYLKGSSGGPLLCPAGHAVGIFRAAVCTRG-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               76; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Query Match 9.1%; Score 87; DB 2; Length 209; Best Local Similarity 26.2%; Pred. No. 0.69; Matches 55; Conservative 18; Mismatches 51; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        17 AQYGLRLLDCLV-----OEEREVHFLISKAAQLVMAT----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   149 GHAVGIFRAAVCTRGVAKAVDFIPVESLET 178
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               --APGFYHQ---PQSVEDLVDFVVARILNT 189
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               angio-associated migratory cell protein - human
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1135 YTKPWETLPREAITHT 1150
                                                                                                                                                                                                                                                                                                                                                                                                                                                                              164 VAKAVDFIPVESLETT 179
```

```
Cyaccession: 199383
R;Beckner, M.E.; Krutzsch, H.C.; Stracke, M.L.; Williams, S.T.; Gallardo, J.A.; Liotta, Cancer Res. 55, 2140-2149, 1995
A;Title: Identification of a new immunoglobulin superfamily protein expressed in blood wa; Reference number: 139383; MUID:95262124; PMID:7743515
A;Reference number: 139383; MUID:95262124; PMID:7743515
A;Reference number: 139383; MUID:95262124; PMID:7743515
A;Reference number: 139383; MUID:97262124; PMID:7743515
A;Reference number: 139383; MUID:9870802; PIDN:AA68889-1; PID:9870803
A;Cross-references: GB:M95627; NID:9870802; PIDN:AA68889-1; PID:9870803
C;Genetics:
A;Genetics:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     54 WIVYHGAGIRITASPKGPVIQMYINVDKDLVGWPAPQGSRSI,----TPCTCGSSDLYI,V 108
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                109 TRHADVIPVRRR---GDSRGS-----LLSPRPISYLKGSSG--GPLLCPA----- 148
C;Date: 06-Sep-1996 #sequence_revision 06-Sep-1996 #text_change 21-Jul-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Query Match
8.6%: Score 82; DB 2; Longth 452;
Best Local Similarity 25.3%; Pred. No. 5;
Matches 42; Conservative 13; Mismatches 47; Indels 64; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         149 ------GHAVGIFR----AAVCTRGVAKAVDFIPVESI, 176
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             296 GSVDCQAKLVSATIGKVVGVFRPETVASQPSLGEGEESESNSVESL 341
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 d
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
```

Search completed: August 30, 2003, 19:20:24 Job time : 17.9789 secs

```
GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.
```

```
(Without alignments)
2506.388 Million cell updates/sec
                                                                                                                                1 MAPITAYAQQTRGLLGCIIT......GVAKAVDFIPVESLETIMRS 182
                                August 30, 2003, 19:20:43; Search time 1764.86 Seconds
 - nucleic search, using frame_plus_p2n model
                                                                                                                                                                                                                                                                        22781392 seqs, 12152238056 residues
                                                                                                                                                                                    0.5
7.0
7.0
                                                                                                                                                                 BLOSUM62
Xgapop 10.0 , Xgapext (
Ygapop 10.0 , Ygapext (
Fgapop 6.0 , Fgapext (
Delop 6.0 , Delext
                                                                                              US-09-965-594-1
953
                                                                                                  Title:
Perfect score:
                                                                                                                                                                      Scoring table:
OM protein
                                                                                                                                    Sequence:
                                                                                                                                                                                                                                                                        Searched:
                                Run on:
```

Minimum DB seq length: 0 Maximum DB seq length: 2000000000

Total number of hits satisfying chosen parameters:

Post-processing: Minimum Match 0% Maximum Match 100% Listing first 45 summaries

Command line parameters:
-MODEL-frame+\_p2n.model -DEV=xlp
-MODEL-frame+\_p2n.model -DEV=xlp
-C-cgn2\_1/USPT0\_spool/US09965594/runat\_29082003\_151919\_28322/app\_query.fasta\_1.2872
-D--Cgn2\_1/USPT0\_spool/US09965594/runat\_29082003\_151919\_28322/app\_query.fasta\_1.2872
-DEST -CPMT-fastap -SUFFIX=rst -MINMATCH-0.1 -LOOPCL-0 -LOOPEXT-0
-UNITS-bits -START-1 -END=-1 -MATRIX=blosum62 -TRANS=buman49 cd1 -LIST-45
-UNITS-bits -START-1 -END=-1 -MATRIX=DIOSUM62 -TRANS=buman49 cd1 -LIST-45
-OUTPRT-pto -NORM-ext -HEAFSIZE=500 -MINIEN=0 -MAXLEN-2000000000
-USBE-US0965594 @CGN 1\_1\_12630\_@runat\_2908303\_151919\_28322 -NCPU-6 -ICPU-3
-NO\_MMAP -LARGEQUERY -NEG\_SCORES-0 -MAIT -DSPBLOK=100 -LONGLOG
-DEV\_IMBOUT-120 -WARN\_TIMBOUT-30 -IHREADS-1 -KGAPDP-10 -KGAPEXT-0.5 -FGAPOP-6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT-0.5 -DELOP=6 -DELEXT=7

em\_gss\_phg:\* em\_gss\_vrl:\* qb\_gssl:\* em\_gss\_pln:\* em\_gss\_vrt:\* em\_gss\_hum:\* em\_gss\_inv:\* ⊕ dss\_fun: \* em\_gss\_mam:• em\_gss\_mus:\* em\_gss\_pro:\* em\_gss\_rod:\* em\_estba:\*
em\_estin:\*
em\_estin:\*
em\_estov:\*
em\_estov:\* em\_estfun:\* em\_estom:\* em\_estro: em\_htc:\* gb\_est1:\* gb\_htc:\* gb\_est3:\* gb\_est3:\* gb\_est5:\* EST:\* Database :

## gp\_gss2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES	

criptid	F30469	Q892487 AGEN	3748 HZ47	M374064 EBma	J001625 BJ0(	J024121 BJ0;	16176 BJ01	F203316 6018	50999 AGEN	0545 Mus	1437 HVSN	3096 Sinc	3244 9630	2274 6018	7233 6016	0860 6024	806 11726	3286 HQ01	9665 CH23	6751 CMD4	8041 6024	9745 AGEN	8830 6030	4657 AGE!	424 AL55	9990 Teta	8584 EBM	4587 AGE	2016	8789	2608	0343	8815	17051	8418	9466	4182	0995	0599	3743	7615 212104	AU12782	2604 BB632604	2790 psMA01	8486 60276049	
di	BF304699	24	37	÷.	97.	7	19	333	909	305	314	9	33	322	72	803	80	332	961	367	980	97	98	346	42	24	982	5.5	220	587	256	503	88	370	84	7.	341	9	305	537	576	82	326	527	984	
ngth DB	984 10	99	15 1	83	15 1	43	54 1	961 1	31 1	141 1	79 1	44 1	01 1	46 1	01 1	58 1	07 9	50 1	39 2	582 1	52 1	283 1	36 1	38	93 9	57 2	470 1	18 1	291	655 1	08 1	411 1	26	32	71 1	98	963 1	29 1	640 1	22 1	39 1	41 9	91	01	44	
& Query Match	11.1	ς,			·.		·.	。	ö	0	。	ö	ö	0	0	·.	ò	·.											٠							٠			٠						•	
COL		•	0	0 (	on (	20.0	σ,	ω.	æ	ω.	97.5	ė.	96	6	95.5	5.	95	σ	94.5	94	94	6	æ.	93.5	<u>ش</u>	93	66	93	σ,	'n	92.2	ς,	76	92	92	3.5	92	92	6	ä	4	91.5	ä	_;	_;	
മജ്	0	7	m ·		S,	ا 0		ထ	σ	0.		c 12	13	c 14	15	c 16	17	18	19	7	7	c 22	7	24	25	26	27	7	c 53	30	c 31	m 1	ا ن	34	35	36	37	m	c 39	4	<b>₹</b>	42	υ •	44	0.45	

## ALIGNMENTS

BF304699 984 bp mRNA linear EST 21-NOV-2000 6018888252F1 NIH\_MGC\_17 Homo sapiens cDNA clone IMAGE:4122276 5', Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. 1 (bases 1 to 984) BF304699.1 GI:11251586 Homo sapiens (human) Homo sapiens mRNA sequence. BF304699 EST. LOCUS RESULT 1 BF304699/c ORGANISM REFERENCE ACCESSION VERSION KEYWORDS SOURCE

~

```
Best Local Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Percent Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    119
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         574
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Ouery Match:
DB:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  source
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           BASE COUNT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Pred. No.:
                                                                                   AUTHORS
TITLE
JOURNAL
COMMENT
                                                           REFERENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    FEATURES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ô
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                    /organism-"Homo sapiens"
//organism-"Homo sapiens"
//mol_type="mrRN"
//db_xref="taxon:9606"
//losuc_type="rhabdomyosarcoma"
/lab_host-"bli0b (phage-resistant)*
//loo_lib="NilLMGC_l7"
//clooe_lib="NilLMGC_l7"
//clooe_lip="NillMGC_l7"
//clooe_lib="NillMGC_l7"
//clooe_lib="Nill"
//clooe_li
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           BQ892487
AGENCOGRE_8417538 Lupski_sympathetic_trunk Homo sapiens cDNA clone IMAGE:6192708 5', mRNA sequence.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        105 LeuTyrLcuValThrArgHisAlaAspVallleProValArgArgArgGlyAspSerArg 124
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              598 ------ACCAGGCACGCAGCGAACATACATGCAAGGAGGACGTGGT---TCCCGC 554
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    140 SerGlyGlyProLeuLcuCysProAlaGlyHisAlaValGlyIlcPheArgAlaAlaVal 159
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            493 CGGCACGCTCCGGTCAGGTGCAGC-----------TTCCAGCGCCCGGGG 455
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             TrpProAlaProGlnGlySerArgSerLeuThr -- - ProCysThrCysGlySerSerAsp 104
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             553 GGGCGCCTCTTGTGGGGAAGACCTCGATGGTGTCCAAGCTCGCGGTGCTGTACTGGAAGT 494
                                                                                                        Email: cgapbs-remail.nih.gov
Tissue Procurement: ArCC
cona Library Preparation: Ling Hong/Rubin Laboratory
cona Library Preparation: Ling Hong/Rubin Laboratory
cona Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
Plate: LLCM1005 row: g column: 13
High quality sequence stop: 646.

1. 984
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 NIH-MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammaiian Gene Collection (MGC)
Unpublished
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       984
333
224
822
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Length:
Matches:
Conservative:
Mismatches:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Indels:
                                                                                         Contact: Robert Strausberg, Ph.D.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    US-09-965-594-1 (1-182) x BF304699 (1-984)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              BQ892487
BQ892487.1 GI:22284501
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     sapiens (human)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    4.34
105.00
45.248
39.298
11.128
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  160 CysThrArgGly 163
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              TGCGCCCGAGGA 443
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      sapiens
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Percent Similarity:
Best Local Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Homo
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             86
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Query Match:
DB:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         VERSION
KEYWORDS
SOURCE
ORGANISM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        DEFINITION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        BASE COUNT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Pred. No.:
                          TITLE
JOURNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ACCESSION
  AUTHORS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       RESULT 2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               BQ892487
                                                                                                                                                                                                                                                                                                                                                                             FEATURES
                                                                                         COMMENT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Q<sub>0</sub>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       QQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
```

•

```
------ArgArgGlyAspSerArgGlySerLeuLeu-- 128
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 484 TGGGATCCATTTTTAATAAGGGTGCTCTGTTAATCATGGCGCCCACGGGCCCGGTGATA 543
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ------CysGlySerSerAsp 104
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          105 LeuTyr-LeuValThr------ ArgHisAlaAspVallleProValArg---- 118
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             -----SerProArgProIleSerTyrLeuLys-------GlySerSerGl 141
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       54 TrpThrValTyrHisGlyAlaGlyThrArgThrIleAlaSerProLysGlyProValIle 73
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             74 GlnMetTyrThrAsnValAspLysAspLeuValGlyTrpProAlaProGlnGlySerArg 93
Eukaryota, Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1199)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ---TGTGCTGCCTGCACACACCCCCATGACCGATGTTGGGGCTTATGTGGAAACGGCGAG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    NIH'MGC http://mgc.nci.nlh.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished
                                                                                                                                                                                                                                                                  þe
                                                                                                                                                   Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Dr. James R. Lupski
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Biosacience Corporation
Clone distribution: MGC clone distribution information can
found through the I.M.A.G.E. Consortium/LLNL at:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             44
6 44
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           /tissue_type="sympathetic trunk"
/dev_stage="adult, 16 yr"
/lab_host="DH10B"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Conservative:
Mismatches:
                                                                                                                                                                                                                                                                                                     http://image.llnl.gov
Plate: LLAM13595 row: c column: 13
High quality sequence start: 57
High quality sequence stop: 394.
Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Matches:
                                                                                                                                                                                                                                                                                                                                                                                                                 1. .1199
/organism="Homo sapiens"
/orl_type="mRNN"
/db_xref="taxon:9606"
/clone="!YAGE:6192708"
/sex="male"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Indels:
                                                                                                                                 Contact: Robert Strausberg, Ph.D.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              US-09-965-594-1 (1-182) x BQ892487 (1-1199)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  94 SerLeuThrProCysThr-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             103.50
37.428
26.458
10.868
```

 $\sim$ 

296

96

9 /

26

114

us-09-965-594-1.rst

8

ò 5 ò

```
Hordeum vulgare subsp. vulgare
Hordeum vulgare subsp. vulgare
Hordeum vulgare subsp. vulgare
Hordeum vulgare subsp. vulgare
Eukaryota; Viridiplanta; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnollophyta; Liliopsida; Poales; Poaceae; Pooldeae;
'Triliceae: Hordeum.
1 (bases 1 to 583)
Hedley, P., Liu, H., Caldwell, D., McCallum, N., Mudie, S., Cardle, L.,
Ramsay, L., Machray, G., Marshall, D.F.M. and Waugh, R.
Development of Barley Transcriptome Resources
Unpublished
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   BM374064 53-JUL-2002 EBma03_SQ003_I16_R maternal, 8 DPA, no treatment, cv Optic, EBma03 Hordeum vulgare subsp. vulgare cDNA clone EBma03_SQ003_I16 5', mRNA
189 GGAGCGGCG------GCATGGTGTTCTTCTGCTTCCTGGTGGACCAGCGGCGGA 236
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           297 TGGCGGACA-------TGGAGACGGCCACCGGCTCTGCTACCTCCTCA 338
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 114 lileProValArgArgArgGlyAspSerArgGlySerLeuLeuSerProArgProlleSe 134
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    999 CCTACCG------CCACTACCGGCTCTACTAGACCCGCGTCGTTGAC 440
                                                                                      ---GlyGluValGlnIleValSer
                                                                                                                                                                                                                                                                                                                                                               77 ThrAsnValAspLysAspLeuValGlyTrpProAlaProGlnGlySerArgSerLeuThr
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              97 ProCysThrCys-----GlySerSerAsp-LeuTyrLeuValThrArgHisAlaAspVa
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 130 CATCAAATACC-TCCGGCGAGAGTTTCAGCGATATTTCTGGAGGAGCTTGGCGCGGGAAGGA
                                                                                                                                                                                                         39 ThrAlaALaGInThrPheLeuAlaThrCyslleAsnGlyValCysTrpThrVal-----
                                                                                                                                                                                                                                                                                                                           57 TyrHisGlyAlaGlyThrArgThrIleAlaSerProLysGlyProValIleGlnMetTyr
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       /clone_lib-"maternal, 8 DPA, no treatment, cv Optic,
EBma03"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Ontact: Waugh R, Marshall DF
Contact: Waugh R, Marshall DF
Contact: Waugh R, Marshall DF
Genome Dynamics/Computational Biology
Scottish Crop Research Institute
Invergowrie, Dundee, DD 2 DA, Scotland, UK
Tel: 00 44 1382 562731
Fax: 00 44 1382 562426
Email: estiscri.ac.uk
All sequence has a phred quality score of 20 or over
Seq primer: M13 reverse.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           /organism="Hordeum vulgare subsp. vulgare"
/mol_type="mRNA"
/cultivar="Optic"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 /clone="EBma03_SQ003_116"
/tissue_type="maternal"
/dev_stage="8 DPA"
/lab_host="DH10B"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       /db_xref="taxon:112509"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         BM374064.2 GI:21941288
                                                                                  28 AsnGlnValGlu---
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          134 rTyrLeu 136
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    441 GCTGATT 447
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        source
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                DEFINITION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            TITLE
JOURNAL
COMMENT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      REFERENCE
AUTHORS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    FEATURES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          BM374064
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                RESULT
                                                                                                                                                                                                                                                     a
                                                                                                                                                                                                                                                                                                                                                                                     a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           C.Y.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Ω
                                                                      ò
                                                                                                                               Oy
Oy
                                                                                                                                                                                                                                                                                                                        8
                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               /dobe_lbb_fy.
/foloe_lbb_fy.
/foloe_lbb_fy.
/foloe_lbb_fy.
/foloe_lob_fy.
/foloe_
                                                                                                                                                                                                                                                                                                                                                                                     CA023748 515 bp mRNA linear EST 23-CCT-2002 HIATTELT HZ HOrdeum vulgare subsp. vulqare CDNA clone HIATTEL7 5-PERME, mRNA sequence. CA023748 CA023748 I GI:24301122
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Eukaryota; Viridiplantae: Štreptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae
AAAACCCCACCTICGGCCCCACCGATGCGCTACCCTTTACAAGCCACCGCCGGG 830
                                                                                                                                            811 CCCCCCCTAACATCTCCTACCCTGCGGCGCGGGGGGGGGAGACGTGGGCGCATACGGGC 870
                                                                                  yGlyProLeuLeuCysProAlaGlyHisAlaValGlyIlePheArgAlaAlaValCysTh 161
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Molecular Markers Group, Department Genbank
Institute of Plant Genetics and Crop Plant Research (IPK)
Corrensstr. 3, 06466, Gatersleben, Germany
Tel: 039482-5595
Fax: 039482-5595
Email: stein@ipk gatersleben.de
Insert Length: 515 Std Error: 0.00
Plate: 47 row: E column: 17
Seg primer: Milzev.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Wobus, U.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               /tissue_type="pericarp"
/dev_stage="0-7 DAP (days after pollination)"
/lab_host="xLl0-Gold"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   /organism-"Hordeum vulgare subsp. vulgare"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Potokina, E. and
                                                                                                                                                                                                         161 rArgGlyValAlaLysAJaValAspPhelleProValGluSer 175
                                                                                                                                                                                                                                                                     871 TCAGGGCGTTTTAAAGCCCCCGGCCTTCGCCGCCGGGGCAAGCA 913
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           515
44
16
16
53
31
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Length:
Matches:
Conservative:
Mismatches:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ; Triticeae; Hordeum.
1 (bases 1 to 515)
Radchuk, V., Zhang, H., Weschke, W.,
Barley ESTs from developing seeds
Unpublished
Contact: Stein Nils
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   /db_xref="GABI:277158"
/db_xref="taxon:112509"
/clone="HZ47E17"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Hordeum vulgare subsp. vulgare
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Hordeum vulgare subsp. vulgare
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             US-09-965-594-1 (1-182) x CA023748 (1-515)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            /mol_type="mRNA"
/cultivar="barke"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       5.6
101.00
41.968
30.778
10.608
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Percent Similarity:
Best Local Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           82
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Alignment Scores:
                                                                                  141
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Query Match:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SOURCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Source
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ..
Q
                                                                                                                                                                                                                                                                                                                                                                                                                     DEFINITION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           BASE COUNT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ACCESSION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              VERSION
KEYWORDS
                                                                                                                                                                                                                                                                                                                                                           CA023748
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         FEATURES
                                                                                                                                                                                                                                                                                                                                  RESULT 3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ORIGIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Score:
```

8 AlaGlnGlnThrArgGlyLeuLeuGlyCysIleIleThrSerLeuThrGlyArgAspLys 27

/note="Vector: pSPORT1; Site\_1: Sal I; Site\_2: Not I;

```
Percent Similarity:
Best Local Similarity:
Query Match:
DB:
                                                                                                                                                                                                                    140
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    BJ02412
                                                                                                                                                                                                                                                                 Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             86
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         111
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              141
                                                                                                                                                                                                                  BASE COUNT
ORIGIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    DEFINITION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ACCESSION
VERSION
KEYWORDS
SOURCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ORGANISM
                                                                                                                                                                                                                                                                                 ..
Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        RESULT 6
BJ024121
                                                 FEATURES
                                                                                                                                                                                                                                                                                              Score:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              q
                                                                                                                                                                                                                                                                                                                                                                                                                  ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                              q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ô
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   d
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ď
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ô
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                BJ001625
BJ001625 MF01SSA cDNA Oryzias latipes cDNA clone MF01SSA025C02 5', mRNA sequence.
Non-normalised library, directionally cloned into pspoRTI. Derived from maternal Lisaue disserted from developing grains. (8 days post anthesis) in glasshouse grown barley plants. Developed as part of the barley transcriptome resources of BBSRC/SERRAD funded cereal IGF (Investigating Gene Punction) profect.

a 176 g 141 t.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ProCysThrCys-----GlySerSerAsp-LeuTyrLeuValThrArgHisAlaAspVa 114
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinoptorygii; Neoptorygii; Teleostei; Butcleostei; Neoteleostei; Acanthomorpha; Acanthopterygii; Percomorpha; Atherinomorpha; Beloniformes; Adrianichthyidae; Oryziinae; Oryzias.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         237
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             114 lileProvalArgArgArgGlyAspSerArgGlySerLeuLeuSerProArgProlleSe 134
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1 (bases 1 to 615)
Kohara,Y., Shin-i,T., Kimura,T., Narita,T., Jindo,T. and Takeda,H.
Medaka EST Project in Takeda's lab
                                                                                                                                                                                                                                                                                                                                                                                    71 CATCAAATACC-TCCGGCGAGAGTTTCAGCGATATITCTGGAGAGCTIGGCGGGAAGGA 129
                                                                                                                                                                                                                                                                                                                                                                                                                                              130 GGAGCGGCG------GCATGGTGTTCTTCTGCTTCCTGGTGGACCAGCGGCGGA 177
                                                                                                                                                                                                                                                                                                                                                                                                                                                                            9.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        96
                                                                                                                                                                                                                                                                                                             11 GCACAGGCGCGCGCGCGCCTAGCTGTTTGCACTGTCCACCACCACCAGGCACCAT 70
                                                                                                                                                                                                                                                                                              AlaGlnGlnThrArgGlyLeuLeuGlyCysIleIleThrSerLeuThrGlyArgAspLys
                                                                                                                                                                                                                                                                                                                                                                                                                                                                              57 TyrHisGlyAlaGlyThrArgThrIleAlaSerProLysGlyProValIleGlnMetTyr
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         178 AGTGCGGAGCAGCAAGCCGCGGCGCGCATCTGCTCGCGGTGCGGCGGCTGCGCCAGCG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        77 ThrAsnValAspLysAspLeuValGlyTrpProAlaProGlnGlySerArgSerLeuThr
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   238 TGGCGGACA-------TGGAGACGGCCACCCGGCTCTGCTACCTCCTCA
                                                                                                                                                                                                                                                                                                                                                        -----GlyGluValGlnIleValSer
                                                                                                                                                                                                                                                                                                                                                                                                                 39 ThrAlaAlaGlnThrPheLeuAlaThrCysIleAsnGlyValCysTrpThrVal----
                                                                                                                                                       583
116
53
31
6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Center For Genetic Resource Information
National Institute of Genetics
1111 Yata, Mishima, Shizuoka 411-8540, Japan
                                                                                                                                                     Length:
Matches:
Conservative:
Mismatches:
Indels:
Gaps:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Oryzias latipes (Japanese medaka)
                                                                                                                                                                                                                                                                (1-583)
                                                                                                                                                                                                                                                                                                                                                       AsnGlnValGlu------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          BJ001625.1 GI:17364516
                                                                                                                                                                                                                                                                US-09-965-594-1 (1-182) x BM374064
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Contact: Tadasu Shin-i
                                                                                                                                                       6.6
101.00
41.968
30.778
10.608
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Oryzias latipes
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          rTyrLeu 136
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    GCTGATT 388
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Unpublished
                                                                                                                                                                                    Percent Similarity:
Best Local Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           BJ001625
                                                                                                                                        Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                       28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         134
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  6
                                                                                                                                                                                                                  Query Match:
DB:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 LOCUS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   RESULT 5
BJ001625/c
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 AUTHORS
TITLE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                JOURNAL
                                                                                                                                                          Pred. No.:
                                                                                             BASE COUNT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   REFERENCE
                                                                                                             ORIGIN
                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                      QQ
                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                qq
                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                          q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ó
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ô
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  oγ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    d
```

```
BJ024121
BJ024121 MF01SSA CDNA Oryzias latipes CDNA clone MF01SSA143D12 3',
MRNA Sequence.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
extinopterygii; Neopterygii; Teleostei; Buteleostei; Neoteleostei;
Acanthomorpha; Acanthopterygii; Percomorpha; Atherinomorpha;
Beloniformes; Adrianichthyidae; Oryzinnae; Oryzias.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       334 CTGCAGGGATGCGGATGTGGCTCTGCT-----TTGGTTCCTCGTCCTCGATCA 284
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               124 ArgGlySerLeuLeuSerProArg------ProlleSerTyrLeuLysGlySerSer 140
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ------TGTTGGAGAACCTATCACAGTTCCTGCTTTAGAGCAACGGCA 410
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        373 --------ACTCGTGGAGGACGAAGACGTCACCCCGGAGCTGTAGG 335
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 223 AGAGGCCGGACAGGCAGTCGGGGGTGAATCTCTGCAGGACGTCTTCACGGCGGATCA 164
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         47 ThrCyslleAsnGlyValCysTrpThrValTyrHisGlyAlaGlyThrArgThrlleAla 66
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           67 SerProLys------GlyProValileGlnMetTyrThrAsnValAspLys 81
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         27 LysAsnGlnValGluGlyGluValGlnIleValSerThrAlaAlaGlnThrPheLeuAla
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           82 AspLeuValGlyTrpProAlaProGlnGlySerArgSerLeuThrPro------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 283 TCTTCCTCACCTGACCTTCCACATCCAGGTGTCGCCAGCGCTGGTCTGACGGGTGATGGG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ------CysThrCysGlySerSerAspLeuTyrLeuValThrArg------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                615
9
50
50
                                                                                                                                                                                                      /db_xref="taxon:8090"
/clone="MPOISSA025C02"
/sex="mixture of female and male"
/tissue_type="whole embryo"
/clove_stage="seqmentation stage 20 -
/clone_lib="MFOISSA cDNA"
a 166 c 165 g 144 t
                                                                                                                                                                                                                                                                                                                 .
                                                                                                                                                                                                                                                                                                                                                                                                                                                                Length:
Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              163 GGAGGACCGACTCGCTGCAGAGCCTCTGCTGCA 131
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   GlyGlyProLeuLeuCysProAlaGlyHisAla 151
                                                                                                       1. .615
/organism="Oryzias latipes"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Oryzias latipes (Japanese medaka)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          US-09-965-594-1 (1-182) x BJ001625 (1-615)
Tel: 81-559-81-6856
Fax: 81-559-81-6855
Email: tshini@genes.nig.ac.jp.
Location/Qualifiers
                                                                                                                                                       /mol_type="mRNA"
/strain="Hd-rR"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       BJ024121.1 GI:17377389
                                                                                                                                                                                                                                                                                                                                                                                                                                                                   11
99.00
33.77%
27.81%
10.39%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Oryzias latipes
```

```
344 GCTCTGCGGCGGCGAGGAGCTCCTGGGCCAGTTGTG---------379
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ------CysThrCysGlySerSerAspLeuTyrLeuValThrArg------ 110
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               469
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   111 ------AisAlaAspValI1eProValArgArgArgGlyAspSer 123
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      529
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       124 ArgGlySerLeuLeuSerProArg-----ProlleSerTyrLeuLysGlySerSer 140
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Oryzias latipes (Japanese medaka)
Oryzias latipes
Oryzias latipes
Dukaryota: Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
Acanthomorpha; Acanthopterygii; Percomorpha; Atherinomorpha;
Beloniformes; Adrianichthyidae; Oryzinae; Oryzias.

1 (bases 1 to 754)
Medaka EST Project in Takeda's lab
Unpublished
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   27 LysasnGlnValGluGlyGluValGlnIleValSerThrAlaAlaGlnThrPheLeuAla 46
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  47 ThrCyslleAsnGlyValCysTrpThrValTyrHisGlyAlaGlyThrArgThrIleAla
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      67 SerProLys------GlyProValIleGlnMetTyrThrAsnValAspLys
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  10 TCTTCCTCACCTGACCTTCCACATCCAGGTGTCGCCAGCGCTGGTCTGACGGGTGATGGG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       82 AspLeuValGlyTrpProAlaProGlnGlySerArgSerLeuThrPro------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             - 25
                                                                                                                                                                                                                                                                                                                                                1...754
/organism="Oryzias latipes"
/mol_type="mRNA"
/strain="Hd-rR"
/db_xref="taxon:8090"
/clone="mrolsSA025202502"
/sex="mixture of female and male"
/tissue_type="whole embryo"
/dev_stage="seqmentation stage 20 -
/clone_llb="MrolsSA adma"
/clone_llb="MrolsSA adma"
/clone_llb="MrolsSA adma"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   754
42
9
50
50
                                                                                                                                                                                                                                                            1111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856
Fax: 81-559-81-6855
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Matches:
Conservative:
Mismatches:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            141 GlyGlyProLeuLeuCysProAlaGlyHisAla 151
                                                                                                                                                                                                            Contact: Tadasu Shin-i
Center For Genetic Resource Information
National Institute of Genetics
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Length:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            US-09-965-594-1 (1-182) x BJ016176 (1-754)
                                                                                                                                                                                                                                                                                                                  Email: tshini@genes.nig.ac.jp.
Location/Qualifiers
 BJ016176.1 GI:17376695
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   14.4
99.00
33.77%
27.81%
10.39%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Percent Similarity:
Best Local Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Alignment Scores
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   302
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Query Match:
                                                                                                                                                                                                                                                                                                                                                         source
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ..
..
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  BASE COUNT
                                                     ORGANISM
                                                                                                                                                        AUTHORS
TITLE
JOURNAL
COMMENT
VERSION
KEYWORDS
SOURCE
                                                                                                                                           REFERENCE
                                                                                                                                                                                                                                                                                                                                        FEATURES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ORIGIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ó
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ó
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           BJ016176 754 bp mRNA linear EST 05-DEC-2001
BJ016176 MF01SSA CDNA Oryzias latipes CDNA clone MF01SSA025C02 3',
mRNA sequence.
BJ016176
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  379
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     418
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             123
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ArgGlySerLeuLeuSerProArg-----ProTleSerTyrLeuLysGlySerSer 140
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              470 TCTTCCTCACCTGACCTTCCACATCCAGGTGTCGCCAGCGCTGGTCTGACGGGTGATGGG 529
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               530 AGAGGCCGCACAGCCAGCAGTCGGGGGTGAATCTCTGCAGGACGTCTTCACGGCGGATCA 589
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            302 -------TGTTGGAGAACCTATCACAGTTCCTGCTTTAGAGCAACGGCA 343
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         242 AAAAATGACGTAGAACCAAAAGACACAGATCCAACCACACATGTTCTGGTTCTACGGGCT 301
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               81
           Kohara,Y., Shin-i,T., Kimura,T., Narita,T., Jindo,T. and Takeda,H.
Medaka EST Project in Takeda's lab
Unpublished
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          47 ThrCyslleAsnGlyValCysTrpThrValTyrHisGlyAlaGlyThrArgThrIleAla 66
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  82 AspleuValGlyTrpProAlaProGlnGlySerArgSerLeuThrPro------ 97
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    27 LysAsnGlnValGluGlyGluValGlnIleValSerThrAlaAlaGlnThrPheLeuAla 46
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 98 ------CysThrCysGlySerSerAspLeuTyrLeuValThrArg------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ----HisAlaAspVallleProValArgArgArgGlyAspSer
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 67 SerProLys ------GlyProVallleGlnMetTyrThrAsnValAspLys
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   380 -----ACTCGTGGAGGACGAAGAGCGTCACCCCGGAGCTGTAGG
                                                                                                                                                                                                                                                                                                                                                         - 254
                                                                                                                                                                                                                                                                                                                                                                                                                                                             643
99
50
50
                                                                                                                                                                                                                                                                           /db xrefw"taxon:8090"
/clone="MF01SSA143D12"
/sex="mixture of female and male"
/fissue_type="whole embryo"
/dev_stage="segmentation stage 20 -
/clone_lib="MF01SSA cDNA"
a 148 c 148 q 176 t
                                                                    Contact: Tadasu Shin-i
Contact: Tadasu Shin-i
Center For Genetic Resource Information
National Institute of Genetics
1111 Yata: Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6855
Fax: 81-559-81-6855
Email: tshini@qenes.nig.ac.jp.
                                                                                                                                                                                                                                                                                                                                                                                                                                                             Length:
Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     141 GlyGlyProLeuLeuCysProAlaGlyHisAla 151
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         /organism-"Oryzias latipes"
/mol_type-"mRNA"
/strain-"Hd-rR"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Gaps:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      US-09-965-594-1 (1-182) x BJ024121 (1-643)
                                                                                                                                                                                                                                                                                                                                                                                                                                                               111.7
99.00
33.77%
27.81%
10.39%
     (bases 1 to 643)
                                                                                                                                                                                                            .643
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Best Local Similarity:
Ouery Match:
DB:
                                                                                                                                                                                                                                                                                                                                                                                            171
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Percent Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             111 ----
                                                                                                                                                                                                                                                                                                                                                                                                                                              Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             LOCUS
                                                                                                                                                                                                                                                                                                                                                                                          BASE COUNT
ORIGIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Pred. No.:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         RESULT 7
BJ016176
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ACCESSION
                                      TITLE
JOURNAL
     REFERENCE
                    AUTHORS
                                                                                                                                                                                                 FEATURES
                                                                          COMMENT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
```

õ 윱 õ g οy ο̈́

9

```
/Organism="miss mussulus"
//mol_type="miss" mussulus"
//db_xref="taxon:10090"
//dlone="lungus" missulus
//dgccattatggcc); Site_1: Sfil
//dgccattatggcc); Site_2: Sfil
//dgccattatggcc); Site_2: Sfil
//dgccattatggcc); Site_2: Sfil
//dgccattatggcc); ChNA made
//dgccattatggcc); Site
//dgccattatggcc); ChNA made
//dgccattatggcc); Site_2: Sfil
//dgccattatggcc); ChNA made
//dgccattatggcc); Site_2: Sfil
//dgccattatggcc); ChNA made
//dgccattatggcc); Site_2: Sfil
//dgccattatggcc); Site_2: Sfil
//dgccattatggcc); Site_2: Sfil
//dgccattatggcc); Site_3: Sfil
//dgccattatggccc, Site_3: Sfil
//dgccattatggccc, Site_3: Sfil
//dgccattatggccc, Site_
                                                                                                                                                                                                                                                              EST 29-APR-2003
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               332 GGAIGTCICATIGATIAIGIGCCTCAAGIGIGCCACCIGCTCACCCACAAGGIICAGIAI 391
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    392 GIGATICAAGGCIAICCCCAAAACAAGAGAGIACAICGGCAGCITITCCIICACICATI 451
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         29 GlnValGluGlyGluValGlnIleValSerThrAlaAlaGlnThrPheLeuAlaThrCys 48
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    16 GlyCysIleIleThr-------SerLeuThrGlyArgAspLysAsn 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              49 IleAsnGlyValCysTrpThrValTyrHisGlyAlaGlyThrArgThrIleAlaSerPro 68
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Eukaryota, Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1 (bases 1 to 1031)
NIH-MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Conpublished
Contact: Robert Strausberg, Ph.D.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Email: ggapbs-r@mail.nih.gov

Tissue Procurement: Dr. Michael Brownstein

Tissue Procurement: Dr. Michael Brownstein

CDNA Library Preparation: Michael Brownstein Laboratory

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)

DNA Sequencing by: Agencourt Bioscience Corporation

Clone distribution: MGC clone distribution information can
found through the I.M.A.G.E. Consortium/LLNL at:

http://lnage.llnl.gov

Plate: NDCM107 row: b column: 11
                                                                                                                                                                                                                   AGENCOURT_13445496 NIH_MGC_177 Mus musculus cDNA clone IMMGE:30316162 5', mRNA sequence.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1031
47
16
59
35
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Length:
Matches:
Conservative:
Mismatches:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    /organism-"Mus musculus"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         =
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      High quality sequence stop: 333.
Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          US-09-965-594-1 (1-182) x CB950999 (1-1031)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Mus musculus (house mouse)
Mus musculus
                                                                                                                                                                                                                                                                                                                                                                                                   CB950999.1 GI:30205777
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            24.3
98.50
40.38%
30.13%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 .1031
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Best Local Similarity:
                                       166 ys 166
                                                                                                             417 AA 416
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   235
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Percent Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Query Match:
DB:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ..
Q
                                                                                                                                                                                                                                                                                             DEFINITION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ORGANISM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               BASE COUNT
                                                                                                                                                                                                                                                                                                                                                                       ACCESSION
                                                                                                                                                                                                                                                                                                                                                                                                           VERSION
KEYWORDS
SOURCE
                                                                                                                                                                                   RESULT 9
CB950999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             PEATURES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ORIGIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Score:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1. 961
//organism="Homo saplens".
//organism="Homo saplens".
//mol_type="mram".
//db_xref="facon:9606"
//db_xref="Thabdomyosarcoma"
//db_host="Thabdomyosarcoma"
//db_host="Dh10B (plage-fesistant)"
//db_host="Dh10B (plage-fesistant)"
//db_host="Dh10B (plage-fesistant)"
//db_host="Dh10B (plage-fesistant)"
//db_host="Dh10B (plage-fesistant)"
//db_host="Dh10B (plage-fesistant)"
//db_host="Dr10B mackel"
//db_host="Dr10B ma
                                                                                                                                        BF203316 961 bp mRNA linear EST 06-NOV-2000 601865914F1 NIH_MGC_17 Homo sapiens cDNA clone IMAGE:4098578 5',
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    106 rLeuValThrArgHisAlaAspValIleProValArgArgArgGlyAspSerArgGlySe 126
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                591 GACTCTGAGTGCGAGGGGGGACAGATGCGGATCGGAAGCGCGGGGGATGCTGATGGGGG 532
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              126 rLeuLeuSer-ProArgProIleSerTyrLeuLysGlySerSerGlyGlyProLeuLeuC 146
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         87 ProAlaProGlnGlySerArgSerLeuThrProCysThrCysGlySerSer-AspLeuTy 106
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ysProAlaGlyHisAlaValGlyIlePheArgAlaAlaValCysThrArgGlyValAlaL 166
                                                                                                                                                                                                                                                                                                                                                                                                                                              Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Email: cgapbs-rémail.nih.gov
Tissue Procurement: AFCC
CDNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
Plate: LLCMSGS row: I column: 03
High quality sequence stop: 637.
Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1 (bases 1 to 961)
NIH-MCC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            961
32
11
32
7
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Matches:
Conservative:
Mismatches:
590 GGAGGACCGACTCGCIGCAGAGCCTCIGCTGCA 622
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Length:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Contact: Robert Strausberg, Ph.D.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Gaps:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      US-09-965-594-1 (1-182) x BF203316 (1-961)
                                                                                                                                                                                                                                                                                         BF203316.1 GI:11096902
                                                                                                                                                                                                                                                                                                                                                                       Homo sapiens (human)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    22.1
98.50
52.44%
39.02%
10.34%
                                                                                                                                                                                                                       mRNA sequence.
BF203316
                                                                                                                                                                                                                                                                                                                                                                                                              Homo sapiens
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Percent Similarity:
Best Local Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         477
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      146
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Query Match:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                source
                                                                                                                                                                                                                                                                                                                            KEYWORDS
SOURCE
ORGANISM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       BASE COUNT
ORIGIN
                                                                                                                                                                                DEFINITION
                                                                                                     BF203316/c
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   AUTHORS
TITLE
JOURNAL
COMMENT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Pred. No.:
                                                                                                                                                                                                                                                          ACCESSION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        REFERENCE
                                                                          RESULT B
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             FEATURES
                                                                                                                                                                                                                                                                                                 VERSION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Score:
                                                                                                                                                    LOCUS
```

õ g õ a à q ð q

þe

polyA_sitc 1141 /note"putative" BASE COUNT 244 a 316 c 353 g 228 t	COMMENT	On Dec 19, 2000 Contact: Wing R Clemson Univers Clemson Univers 100 Jordan Hall
Alignment Scores: 27.7 Length: 1141 Score: 98.50 Marches: 51 Percent Similarity: 34.83% Conservative: 11 Best Local Similarity: 28.65% Mismatches: 59 Query Match: 10.34% Indels: 57 DB: 10.34% Gaps: 10	FEATURES	Tel: 864 656 722 Fax: 864 656 422 Email: rwing@cl Total hq bases. Seq primer: AAr' High quality see
US-09-965-594-1 (1-182) x AKO8U545 (1-1141)	source	
Qy     13 GlyLeuLeuGlyCysIleIleThrSerLeuThrGlyArgAspLysAsnGlnValGluGly 32       11   1   1   1   1   1   1   1   1   1		/mol_t. /culti: /db_xr /clone
Qy     33 GluValGinileValSerThrAlaAlaGinThrPheLeuAlaThrCysIleAsnGiyVal 52       Db     306 GAAGCGCAGCTGGACGCGCGCGCAG		/tissu/lab_h /lab_h /clone HVCDNA
Qy 53 CysTrpThrValTyrHisGlyAlaGlyThrArgThrIleAlaSerProLysGlyProVal 72		/note= Seeds condit with w
Qy     73 IleGlnMetTyrThrAsnValAspLysAspLeuValGlyTrpProAlaProGlnGlySer 92       Db     375 GTGCCTCCTCAAGGAGGG 392		crysta incuba total primar
Oy 93 ArgSerLeuThrProCysThrCysGlySerSerAspLeuTyrLeuValThrArgHisAla 112		in viv These the Un Fenton
Oy 113 AspVall1eProValArgArgArgGlyAspSerArgGlySerLeuLeuSerProArg 131		Univer Frisch sequen (Wing,
Oy 132 Prolle		The se and co above.
Qy     146 Cysproala GlyHisalavalGlyIlePhcargalavalCysThrargGlyVal 164       IIIII   II :::IIIII::     IIIII       bb     549 GCCCCAGCGAAGGTCTGTAGGGTTG GCCCTGTAAGTGTGTGAGTGCGAGAC 605		http:/ this c see Cl Geneti
Oy 165 AlabysalaValAspPhelleProValGluSerIeuGluThrThrMet 180	BASE COUNT ORIGIN	barley (http: 147 a 21
BEG31437  BEG31437  LOCUS  BF631437  LOCUS  DEFINITION HVSMED0015P05f Hordoum vulgare seedling shoot EST library HYCODNA002 (Dehydration Stress) Hordeum vulgare subsp. vulgare cDNA	Alignment Scores: Pred: No.: Score: Percent Similarity Best Local Similar	Alignment Scores: 21 Score: 97. Percent Similarity: 39. Percent Similarity: 30.
NOI NO	DB: DB: US-09-965-	4-1 (1-182)
WS.	A A	35 GlnIleValSerT    :::   CAGCTGCCAACGG
REFERENCE 1 (bases 1 to 779) AUTHORS Wing,R., Close,T.J., Kleinhofs.A., Wise,R., Begum,D., Frisch,D., Yu ,Y., Henry,D., Palmer,M., Rambo,I., Simmons,J., Choi,D.W., Fenton	oy ag	55 ThrValTyrHisG        :::   49 ACCCGTGGAGGG
,R.D., Oates,R. and Main,D. TITLE Development of a genetically and physically anchored EST resource for barley genomics: Morex drought-stressed seedling shoot CDNA	Λ̈́O	74
llbrary JOURNAL Unpublished	<b>q</b> 0	109 CCCTGTTCTGCCA

```
clone_lib_"Hordeau vulgare seedling shoot EST library vcDNA002 (Dehydration stress)*

**NEDNA002 (Dehydration stress)**

**NEDNA002 Sterilized then germinated under axenic enditions in the dark at room temperature on filter paper ith water, nystain axenic deformed in covered contains at 100 RH for 24 hr. Shoots were then harvested, otal IRA was prepared, Dely(A) RNA was putified, one rimary unamplified CDNA library was made, 600000 pfu were rivo excised to give pBluescript SK(-) CDNA phagemids, hese steps were performed in the TJ Close laboratory at the University of Callfornia, Riverside (Choi, Close, enton). Phagemids were plated and picked at the Clemson niversity Genomics Institute (CUGI) (Begum, Palmer, risch, Atkins and Wing), Plasmid DNA preparations, DNA equencing and sequence analysis were performed at CUGI wing, Yu, Frisch, Henry, Simmons, Oates, Rambo, Main). The sequence has been trimmed to remove vector sequence have been trimmed to remove vector sequence.

**Part of the contains a minimum of 100 bases of phred value 20 or bover.**
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    AGGTGGCCGCGGTGGACGGCGCGCGGGAGGGAAGCGGCTCGTGGTCG 168
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   //www.genome.clemson.edu/projects/barley. To order clone see http://www.genome.clemson.edu/orders Also close TJ, Wing R, Kleinhofs A, Wise R (2001) ically and physically anchored EST resources for sy genomics. Barley Genetics Newsletter 11:29-30. sty genomics. Barley Genetics Newsletter 11:29-30. sty genomics. Ju go Control of the TJ (10 to TJ) in go TO to TJ (10 to TJ) is the TJ (10 to TJ) in go TO to TJ (10 to TJ) is the TJ (10 to TJ) in go TJ (10 to TJ) in go
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      GCGCCCCCCCGCGAACAACTACAGGACGGTGC------TGCTGG 48
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    GlyAlaGlyThrArgThrIleAlaSerProLysGlyProVallle--- 73
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ThralaAlaGlnThrPheLeuAlaThrCysIleAsnGlyValCysTrp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 -----GlnMetTyrThrAsnValAspLysAspLeuValGlyTrpPro
0 this sequence version replaced gi:11895595.
RA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         nism-"Hordeum vulgare subsp. vulgare"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    779
47
14
63
30
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Length:
Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                       Clemson, SC 29634, USA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           tref="taxon:112509"
ne="HVSMED0015P05f"
sue_type="Seedling shoot"
.host="TJC121"
                                                                     sity Genomics Institute sity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Gaps:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  TTAACCCTCACTAAAGGG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     equence stop: 588.
.ion/Qualifiers
79
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          nce analysis see
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               BF631437 (1-779)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              type="mRNA"
ivar="Morex"
                                                                                                                                                                                                                                                                                                                                                                                                     lemson.edu
                                                                                                                                                                                                                                              1, c
288
```

ATC	563 CCCGGTTCAAGAAGCTGCTGATGCTGGGCGAGCGGGCCCTCTTGCGC 513	AsnValAspLysAspLeuValGlyTrpProAlaProGlnGlySerArgSerLeuThrPro	512GIZECCCC 498 98 CysThrCysGIYSerSerAspLeuTyrLeuValThrArgHisAlaAsp 113			134 SerTyrLeuLysGlySerSerGlyGlyProLcuCuseuCysProAlaGlyHisAlaValGly 153	419 GCATICTGCGGGGCAACTGGAAGTGCCTTACCCTGCCGTATTCATCAACTTCAACT 363 154 IlepharqalaalayalCvsThrargGlyValAlaLaLvsAlaValasp 169		170 PhelleProval 173 :::        308 GGIGTTCCIGTC 297	LT 13	NOILIN	ACCESSION BE863244 GI:12253388	_		RS	TITLE Analyses of the Chlamydomonas reinhardtii Genome: A Model, Unicellular System for Analyzing Gene Function and Regulation in Vascular Plants; project phase 3	JOKWAL Unpullated COMMENT CONTEST Hauser  COMMENT POWN 1000  Duke University	Durham, NC 27708-1000 Tel: 919 613 8159 Fax: 919 613 8179 Eax: 919 613 6170 Email: chauser@duke.edu. FEATURES 1_coration/Qualifiers	10/ DE/	/db_xref="taxon:3055" 1f //db_xref="taxon:3055" //db //db //db //db //db //db //db //d	normalized, Lambda Zap II" /note="Vector: pBluescript II SK-; Site_1: EcoRI; Site_2: Xhoi; This library, constructed by John Davies and Jeffrey Manager Cambines Danse from Co-1600 calls aroun to	mid-log phase in TAP-N (30 min, lhr, 4hr), TAP-S (30 min, lhr, 4hr), TAP-P (4hr, 12hr, 24hr), NO3 to NH4 (30min, lhr, 4hr) and NH4 to NO3 (30min, lhr, 4hr). PolyA mRNA was	purified from each sample, pooled and cDNA synthesized. The cDNA was directionally cloned into lambda Zap II (Stratagene) in the EcoRI (5') and XhoRI (3') sites.
d yo	qa	δ	qa is	<b>q</b> &	qq	٥y	සු ර	QQ QQ	Qy	RESULT 1	LOCUS	ACCES	KEYWORDS SOURCE ORGANIE	12 12 12 12 12	<b>A</b> UT	TIT C	DO CO <b>NN</b>	PEAT					
Qy 88 AlaproGlnGlyScrArgSer	Qy 104 -AspLeuTyrLeuValThrArgHisAlaAspValIleProValArgAr 119		Oy 119 gArgGlyAspSerArgGlySerLeuLeuSerProArgProlleSerTyrLeuLysGlySe 139	Oy 139 rSerGly-GlyProLeureuCySProAla	151 laValGlyIlePheArqAlaAlaValCysThrArg 162	DD 408 GCGGAGGCGCGACGACGCGGGGCTCTGCGT 442	RESULT 12 CNSO9D4S/C LOCUS CNSO9D4S 844 bp mRNA linear HTC 08-JAN-2003	ITION	z		URGANISH ANDPILLES SUMULAE Enkaryota: Metazoa: Arthropoda: Hexapoda: Insecta; Pterygota; Neoptera; Endopterygota: Diptera: Nematocera: Culicoldea:	REFERENCE 1 (bases 1 to 844) AUTHORS Genescope.		FEATURES WWW genoscope.cns.fr)	) }	/strain="6-9" /db_xref="taxon:7165" /clone="Fk0AG10Cc12"	/Plasmid~"PMELBS-FE." /note-"end : 3-PRIME. BASE COUNT 197 a 220 c 213 g 214 E	Alignment Scores: 29 Longth: 844  Pred. No.: 29 Matches: 43  Score: 96.50 Matches: 43  Score: 35.87 Conservative: 23  Boat Long Similarity: 23.37 Mismatches: 63	10.13% Indels: 11 Gaps:	1 (1-182) x CNSO9D4S (1-844)	Qy     10 GlnThrArgGlyLeuLeuGlyCysIleIleThrSerLeuThrGlyArgAspLysAsnGln 29       1 ::       ::       ::	Qy       30 ValGluGlyGluValGlnileValSerThrAlaAlaGinThrPheLeuAlaThr 47         ::       ::       ::	Oy 48 CysileAsnClyValCysTrpThrValTyrHisGlyAlaClyThrArgThrIle 65

```
Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              EST.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        555
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        438
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       90
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         162
                                                                                                                                                                                                                                                                                                                         Query Match:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         LOCUS
DEFINITION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         VERSION
KEYWORDS
SOURCE
ORGANISM
                                                                                                                                                                                     BASE COUNT
ORIGIN
                                                                                                                                                                                                                                                          ..
9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        RESULT 15
BF307233
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            REFERENCE
AUTHORS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           TITLE
JOURNAL
COMMENT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ACCESSION
                                                                                                                                                                                                                                        Alignment
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ď
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        q
                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ç
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      BF182274 846 bp mRNA linear EST 31-OCT-2000 601804028F1 NCI_CGAP_Mam5 Mus musculus cDNA clone IMAGE:4035102 5',
pHluescript II SK- plasmids were excised from the lambda ZAP clones by superinfection with ExAssist (Stratagenc) hadge. The library was normalized using method 4 described in Bonaldo et al. 95 Genome Research 6: 791-866.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Contact: Robert Strausberg, Ph.D.

Email: cgapbs-r@mail.nih.gov
   Tissue Procurement: Lothar Hennighausen Ph.D., Robin Humphreys
   CDNA Library Preparation: Life Technologies, Inc.
   CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
   DNA Sequencing by: Incyte Genomics, Inc.
   Clone distribution: MGC clone distribution information can be
   found through the I.M.A.G.E. Consortium/LLNL at:
   http://image.lln.gov
   Plate: LLAM9308 row: g column: 07
   High quality sequence stop: 696.
   Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        124
                                                                                                                                                                                                                                                                                                             :::III
:::III
CACCACCATACCCTTGCTCTCAGTTGCTCACCAAAATTATGCCCATACGGGCCACTA 230
                                                                                                                                                                                                                                                                                                                                                                                       284
                                                                                                                                                                                                                                                                                                                                                                                                                     93 ArgSerLeuThrProCysThrCysGlySerSerAspLeuTyrLeuValThrArgHisAla 112
                                                                                                                                                                                                                                                                                                                                                                                                                                            404
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     405 GAAGGGACGGAATTCCGTTTGCGGAAAAGTGAGCGCCCCAAGGTCTGACCAAGTGCTA 464
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Mus musculus sukaryota: Chordata, Craniala: Vertebrata, Euteleostomi: Bukaryota: Metazoa: Chordata, Czaniala: Vertebrata, Eutheria: Rodentia: Sciurograthi; Muridae; Murinae; Musinae; Lo 846)
1 (bases 1 to 846)
11 (bases 1 to 846)
Nul+McC http://mgc.nci.nih.gov/
National Institutes of Health, Mammalian Gone Collection (MCC)
                                                                                                                                                                                                                                                                                                                                                     73 IleGlnMetTyrThrAsnValAspLysAspLeuValGlyTrpProAlaProGlnGlySer 92
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    345 GAGATCATTTGCATGTGGCTTTAGTCACCCCAAGAAGAGCCCTGGGAGTGGGCATTTATAA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     TyrHisGlyAlaGiyThrArgThrileAlaScrProLys------GlyProVal
                                                                                                                                                                                                                                                                                                                                                                                     231 ACAAAGTTACATACACGG -----AAGGACCAGCGCGCTTGGCCACCCCCTTGGAGCCG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        -----ProvalArgArgArgGlyAspSerArg----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     137 LysGlySerSerGlyGlyProLeuLeuCysProAlaGly 149
                                                                                                                                    701
32
14
45
45
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CICCAAGGCAGCAAIGGGAGCCITTCGCGGTGIGCGGGT 503
                                                                                                                                  Length:
Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           /organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
                                                                                                                                                                                                                                                    US-09-965-594-1 (1-182) x BF863244 (1-701)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Mus musculus (house mouse)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     BF182274.1 GI:11050416
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      113 Aspvalile-----
                                                                                                                                  25.4
96.00
40.718
28.328
10.078
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1. .846
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    mRNA sequence.
BF182274
                                                                                                                                                                                 Best Local Similarity:
                                                                    Ø
                                                                                                                                                                   Percent Similarity:
                                                                                                                      Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       EST.
                                                                                                                                                                                                                                                                                                                                                                                                                                                    285
                                                                                                                                                                                                     Query Match:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      LOCUS
DEFINITION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ORGANISM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    BF182274/C
                                                                                                                                     ..
Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
                                                                  BASE COUNT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ACCESSION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   VERSION
KEYWORDS
SOURCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  RESULT 14
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             FEATURES
                                                                                  ORIGIN
                                                                                                                                  Pred.
                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                QQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                    a
                                                                                                                                                                                                                                                                                                                                                                                  Ω
                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     6
                                                                                                                                                                                                                                                                                      ð
                                                                                                                                                                                                                                                                                                                                                    ò
```

```
BF307233 901 bp mRNA linear EST 21-NOV-2000 001891502F1 NIH_MGC_17 Homo sapiens cDNA clone IMAGE:4137145 5', mRNA sequence.
BF307233
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            128 uSerProArgProlleSerTyrLeuLysGlySerSerGlyGlyProLeu---LeuCysPr 147
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         /clone="IMAGE:4035102"
/tissue_type="tumor, gross tissue"
/dev_stagg="type="tumor, gross tissue"
/dev_stagg="type="tumor, gross tissue"
/lab_host="bh108"
/clone_lib="NCI_CGAP_Mam5"
/note="Organ: mammary; Vector: pCMV-SPORT6; Site_1: Sall;
Site_2: Not!; cloned unidirectionally. Primer: Oligo dr.
Library constructed by Life Technologies. Investigators
providing samples: Lothar Hennighausen/Robin Humphreys,
NIH
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            147 oAlaGlyHis---AlaValGlyIlePheArg------AlaAlaValCysThrAr 162
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               757 GGTTCTGCTACCAGAAACAGCCTGGATGAAGAAGGACCA-------CATCCTTC 710
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          709 GGITCTICAGICCCAAGIGGGCIGGAGAAAGIGAACACAGAGAAAAGAGGACGGICCCICA 650
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            nGlySerArg-----SerLeuThrProCysThrCysGlySerSerAspLeuTyrLeuVa 108
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Homo Sapiens
Homo Sapiens
Eukaryota: Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalla; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 901)
NIH-WGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          GlyAlaGlyThrArg-ThrIleAlaSerProLysGlyProValIleGlnMetTyrThrAs
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          108 1ThrargHisAlaAspVallleProValArgArgArgGlyAspSerArgGlySerLeuLe
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          609 -ACTAGACACACCT -- GTAATCCAGGAGGAACGCNTGGAGGAACAGAGGGACTCC --- CT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         cDNA Library Preparation: Ling Hong/Rubin Laboratory cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
                                                                                                                                                                                                                                                                             1 others
                                                                                                                                                                                                                                                                                                                                                                               846
446
113
36
31
9
                                                                                                                                                                                                                                                                                                                                                                                                     Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                             210 t
                                                                                                                                                                                                                                                                                                                                                                                  Length:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Gaps:
/db_xref-"taxon:10090"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          US-09-965-594-1 (1-182) x BF182274 (1-846)
                                                                                                                                                                                                                                                                          241 9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    BF307233.1 GI:11254342
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Homo sapiens (human)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               gGlyValAlaLys 166
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            TGGGATGACTGAG 426
                                                                                                                                                                                                                                                                          218 c
                                                                                                                                                                                                                                                                                                                                                                             32.5
96.00
47.20%
36.80%
                                                                                                                                                                                                                                                                                                                                                                                                                               Percent Similarity:
Best Local Similarity:
                                                                                                                                                                                                                                                                          176
```

```
DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be cloud through the I.M.A.G.E. Consortium/LLKI. at: image.llnl.gov plate: LLCM1044 row: c column: 02
High quality sequence start: 6
High quality sequence stop: 684.

Location/Cualifiers

Amol_type="mRNA" e846"

/do_xrof="taxon:9606"
/do_x
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   86 pProAlaProGlnGlySerArgSerLeuThrProCysThrCysGlySerSerAspLeuTy 106
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   106 rLeuValThrArgHisAlaAspValIleProValArgArgArgGlyAspSerArgGlySe 126
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  -----GTGAGCCTTCGTCAGGGGGGGGTTGGGCGGGGTTC 766
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  rLeuLeuSerProArgProIleSerTyrLeuLysGlySerScrGlyGlyProLeuLeuCy 146
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       66 aSerProLysGlyProValIleGlnMetTyrThrAsnValAsplysAspLeuValGlyTr 86
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        sProAlaGlyHisAlaValGly-----IlePheArgAlaAlaValCys 150
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       901
37
8
8
45
5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Longth:
Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Gaps:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            US-09-965-594-1 (1-182) x BF307233 (1-901)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      39.5
95.50
38.46%
31.62%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Percent Similarity:
Best Local Similarity:
Query Match:
DB:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           732
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     146
                                                                                                                                                                                                                              source
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Pred. No.:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 BASE COUNT
                                                                                                                                                                                               FEATURES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ORIGIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ô
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
```

Search completed: August 31, 2003, 04:27:18 Job time : 1773.86 secs

```
GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.
```

OM protein - protein search, using sw model

August 30, 2003, 17:42:58; Search time 44.1697 Seconds Run on:

(without alignments)
700.745 Million ceil updates/sec

US-09-965-594-12

1 MKKKGSVVIVGRIVLNGAYA......VAKAVDFIPVESLETIMRSP 195 Perfect score: Sequence:

BLOSUM62 Gapop 10.0 , Gapext 0.5 Scoring table:

1107863 seqs, 158726573 residues Searched:

1107863 Total number of hits satisfying chosen parameters:

length: 0 length: 2000000000 Minimum DB seq Maximum DB seq

Post-processing: Minimum Match 08 Maximum Match 1008 Listing first 45 summaries

Database :

| SIDSI/gcddata/geneseq/geneseqp-embl/AA1980.DAT:
| SIDSI/gcdata/geneseq/geneseqp-embl/AA1981.DAT:
| SIDSI/gcddata/geneseqy-embl/AA1981.DAT:
| SIDSI/gcddata/geneseqy-embl/AA1983.DAT:
| SIDSI/gcgdata/geneseqy-embl/AA1983.DAT:
| SIDSI/gcgdata/geneseqy-embl/AA1985.DAT:
| SIDSI/gcgdata/geneseqy-embl/AA1985.DAT:
| SIDSI/gcgdata/geneseqy-embl/AA1985.DAT:
| SIDSI/gcgdata/geneseqy-embl/AA1986.DAT:
| SIDSI/gcgdata/geneseqy-embl/AA1989.DAT:
| SIDSI/gcgdata/geneseqy-embl/AA1989.DAT:
| SIDSI/gcgdata/geneseqy-embl/AA1991.DAT:
| SIDSI/gcgdata/geneseqy-embl/AA1997.DAT:
| SIDSI/gcgdata/geneseqy-embl/AA1998.DAT:
| SIDSI/gcgdata/geneseqy-embl/AA1998.DAT: /SIDS1/gogdata/geneseq/geneseqp·emb1/AA1999.DAT:\*/SIDS1/gogdata/geneseq/geneseqp·emb1/AA2000.DAT:\*/SIDS1/gogdata/geneseqy/geneseqp·cmb1/AA2001.DAT:\* /SIDS1/gcgdata/geneseq/genesegp-embl/AA2003.DAT:\* /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2002.DAT: A\_Geneseq\_19Jun03:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Description	Hepatitis C virus								
	ID		AAB15221	AAB15212	AAB15222	AAB15226	AAB15223	AAB15224	AAB15225	AAY44728
	DB	21	21	21	21	21	21	21	21	21
	re Match Length DB	195	197	195	197	197	197	197	197	191
% Ouery	Match	100.0	97.7	96.2	1.96	93.9	93.1	92.0	91.0	88.1
	Score	1021	866	885	981	959	951	939	929	899.5
Result	Q	-	7	m	4	S	9	7	œ	6

Modified hepatitis C virus (HCV) NS3 protease comprising at least 1

WPI; 2000-465976/40. N-PSDB; AAA73329.

HCV-1 polyprotein.  HCV genomic amino.  Sequence encoded 1  Protein sequence of Peptide encoded 1  Sequence encoded 1  Sequence encoded 1  Protein encoded by Composite hepatiti Amino acid sequence Polyprotein encode Compiled HCV sequence Compiled HCV sequence Hepatitis C virus HCV polyprotein.  HCV polyprotein in HCV polyprotein.  HCV polyprotein.  HCV polyprotein in Hepatitis C virus Hepatitis C virus HCV polyprotein.  HCV polyprotein in Hepatitis C virus HCV polyprotein.  HCV polyprotein in Hepatitis C virus HCV-1 NS3/Ad mutan Hepatitis C virus HCV-1 NS3/Ad a confo Hepatitis C virus HCV-1 NS3/Ad a confo Hepatitis C virus	HCV NS3 protein. HCV NS4A-NS3 compl HCV POIVPOPTIOE 1. HCV amino acid seq Hepatitis C virus
AAR34009 AAR40120 AAP90158 AAP90164 AAP90164 AAP92080 AAP92080 AAP18540 AAR31621 AAR31637 AAR31637 AAR56699	AAN93482 AAY24943 AAR28582 AAR28582 AAR28123 AAE1847 AAE1839 AAE21839
440000011212222222222222222222222222222	23333333333333333333333333333333333333
2816 11766 11766 11766 12261 22436 22436 22436 22436 22436 22436 22436 22436 23011 3011 101 101 101 101 101 101 101 10	24435 24435 24436 2472 6332 686 686
9 9 9 8 9 7 7 7 9 9 9 9 9 9 9 9 9 9 9 9	
о о о о о о о о о о о о о о о о о о о	. 9 9 9 9 9 9 9 9 9 9 9 9 9 9 9 9 9 9 9
01111111111111111111111111111111111111	0 W W W W 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4

## ALIGNMENTS

Hepatitis, NS3 protease, viral replication; chronic liver disease; liver failure; liver cancer; mutant; mutein. Goldfarb V; Hepatitis C virus NS4A-NS3 fusion protease #2. Zhang Y, AAB15220 standard; protein; 195 AA (BRIM ) BRISTOL-MYERS SQUIBB CO. 06-JAN-2000; 2000WO-US00345. 99US-0115271, Wittekind M, Weinheimer S, 19-DEC-2000 (first entry) Hepatitis C virus. Synthetic. WO200040707-A1. 08-JAN-1999; 13-JUL-2000 AAB15220; RESULT 1 

```
;;
                                                                                                                                                                                              fusion proteins which can be used to identify inhibitors of this type, as well as enabling structural studies of the protease and protease:inhibitor complexes. This sequence contains the alpha-helix0-1
                                                                                                                                                                                                                                                                                                                                                                                                            INGVCWTVYHGAGIRTIASPKGPVIQMYTNVDKDLVGWPAPQGSRSLTPCTCGSSDLYLV 120
                                                                                                                                                                                                                                                                                                                                                                                                                            TRHADVIPVRRKGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVCTRGVAKAV 180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                The present sequence is a mutated version of a fusion protein created using the Hepatitis C virus (HCV) NS3 and NS4A protease enzymes. These proteins are both essential for the replication of the virus, acting to cleave its replicative proteins from the polyprotein produced from the HCV genome. Inhibitors of the two proteins should be effective as antiviral treamments of HCV infection. This is useful as HCV can load to chook it invention concerns a number of NS3 mutants and NS3-NS4A
                                                                                                                                                                                                                                                                                                                                                        9
                                                                                                                                                                                                                                                                                                                                                                                   9
substitution of a hydrophobic alpha-helix 0 amino acid to a hydrophilic amino acid, useful for screening inhibitors that may treat hepatitis C
                                                                                                                                                                                                                                                                                                                                                                       1 MKKKGSVVIVGRIVLNGAYAQQTRGEEGCQETSQTGKDKNQVEGEVQIVSTAAQTFLATC
                                                                                                                                                                                                                                                                                                                              Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     protease; viral replication; chronic liver disease; liver cancer; mutant; mutein.
                                                                                                                                                                                                                                                                                                                           0
                                                                                                                                                                                                                                                                                               Query Match 100.0%; Score 1021: DB 21; Length 195; Best Local Similarity 100.0%; Pred. No. 2.7e-98; Matches 195; Conservative 0; Mismatches 0; Indels 0;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Goldfarb V;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Hepatitis C virus NS4A-NS3 fusion protease #3.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Zhang Y,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Ź
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ဗ
                                                        Claim 23; Fig 12; 66pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         AAB15221 standard; protein; 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (BRIM ) BRISTOL-MYERS SQUIRE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          DFIPVESLETTMRSP 195
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        DFIPVESLETTMRSP 195
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Wittekind M, Weinheimer S,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               2000WO-0500345
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          99US-0115271
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (first cntry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              WPI; 2000-465976/40.
                                                                                                                                                                                                                                                                    195 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            N-PSDB; AAA73330.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    liver failure:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       WO200040707-A1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               06-JAN-2000;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          08-JAN-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 19-DEC-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   13-JUL-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Hepatitis:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Synthetic
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    AAB15221;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         181
                                                                                                                                                                                                                                                                      Sequence
                                                                                                                                                                                                                                                                                                                                                                                                               61
                                                                                                                                                                                                                                                                                                                                                                                                                                        61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  181
                                                                                                                                                                                                                                            variant
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             RESULT 2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ga
                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                        ö
                                                                                                                                                                                                                                                                                                                                                                                                             à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
```

```
The present sequence is a mutated version of a fusion protein created using the Hepatitis C virus (HCV) NS3 and NS4A protease enzymes. These proteins are both essential for the replication of the virus, acting to cleave its replicative proteins from the polyprotein produced from the antiviral treatments of the two proteins should be effective as antiviral treatments of the two proteins should be effective as antiviral treatments of the involutions of the section. This is useful as HCV an lead to chronic liver disease such as cirrhosis, liver failure and liver cancer. The present invention concerns a number of NS3 mutants and NS3-NS4A fusion proteins which can be used to identify inhibitors of this type, as well as enabling structural studies of the protease and protease:inhibitor complexes. This sequence contains the alpha-helix0-1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   119 LVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVCTRGVAK 178
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1 MKKKGSVVIVGRIVLNG---AYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVSTAAQTFLA 58
Modified hepatitis C virus (HCV) NS3 protease comprising at least 1 substitution of a hydrophobic alpha-helix 0 amino acid to a hydrophilic amino acid, useful for screening inhibitors that may treat hepatitis C
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1 MKKKGSVVIVGRINLSGDTAYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVSTAAQTFLA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  59 TCINGVCWTVYHGAGTRTIASPKGPVIOMYTNVDKDLVGWPAPQGSRSLTPCTGSSDLY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                61 TCINGVCWTVYHGAGTRTIASPKGPVIQMYTNVDKDLVGWPAPQGSRSLTPCTCGSSDLY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  depatitis; NG3 protease; viral replication; chronic liver disease;
liver failure; liver cancer.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            .;
;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Length 197;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Goldfarb V;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Score 998; DB 21;
Pred. No. 7e-96;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Repatitis C virus NS4A-NS3 fusion protease #1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ۲
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 2hang
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Ź
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        AAB15212 standard; protein; 195
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                (BRIM ) BRISTOL-MYERS SQUIBB CO.
                                                                                                                                                                    English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         179 AVDFIPVESLETTMRSP 195
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  97.78;
98.08;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Weinheimer S,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              36-JAN-2000; 2000WO-US00345
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  99US-0115271
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Matches 193; Conservative
                                                                                                                                                                    23; Fig 13; 66pp;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               WPI; 2000-465976/40.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          197 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Repatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 N-PSDB; AAA73328
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               40200040707-A1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Wittekind M,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               08-JAN-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       19-DEC-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                13-JUL-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Synthetic.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Sednence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         AAB15212;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                rariant.
                                                                                                                                                                        Claim
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      AAB15212
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   RESULT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      DER WEEK STANDER STAND
      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         đ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              9
```

```
120
                                                                                                                                                                                                                                                                                                                                                                                                                                        TRHADVIPVRRRGDSRGSI,LSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVCTRGVAKAV 180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               60
                                                                                                                                                                                                                                                                                                                                                                                           9
                                                                                                        The present sequence is a fusion protein created using the Hepatitis C virus (HCV) NS3 and NS4A protease enzymes. These proteins are both essential for the replication of the virus, acting to cleave its replicative proteins from the polyprotein produced from the HCV genome. Inhibitors of the two proteins should be effective as antivizal treatments of HCV infection. This is useful as HCV can lead to chronic liver disease such as cirrhosis, liver failure and liver cancer. The present invention concerns a number of NS3 mutants and NS3-NS4A fusion proteins which can be used to identify inhibitors of this type, as well as enabling structural studies of the protease and protease:inhibitor
             least 1
hydrophilic
                                                                                                                                                                                                                                                                                                                                                                                 1 MKKKGSVVIVGRIVLNGAYAQQTRGEEGCQETSQTGRDKNOVEGEVQIVSTAAQTF1.ATC
                                                                                                                                                                                                                                                                                                                                                                                                                         INGVCWTVYHGAGIRTIASPKGPVIQMYTNVDKDLVGWPAPQGSRSLTPCTCGSSDLYLV
             Modified hepatitis C virus (HCV) NS3 protease comprising at least 1 substitution of a hydrophobic alpha-helix 0 amino acid to a hydrophil amino acid, useful for screening inhibitors that may treat hepatitis
                                                                                                                                                                                                                                                                                                                                        Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Hepatitis; NS3 protease; viral replication; chronic liver disease;
liver failure; liver cancer; mutant; mutein.
                                                                                                                                                                                                                                                                                                                                       Ċ.
                                                                                                                                                                                                                                                                                                           Length 195;
                                                                                                                                                                                                                                                                                                                                     Indels
                                                                                                                                                                                                                                                                                                         Score 982; DB 21;
Pred. No. 3.2e-94;
0; Mismatches 5;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Goldfarb
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Hepatitis C virus NS4A-NS3 fusion protease #4.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Shang Y,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Ş
                                                                                 Example 2; Fig 10; 65pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       S
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        AAB15222 standard; protein; 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       (BRIM ) BRISTOL-MYERS SQUIBB
                                                                                                                                                                                                                                                                                                         Match 96.2%;
Local Similarity 97.4%;
nes 190; Conservative (
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        DFIPVESLETTMRSP 195
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       DFIPVESLETTMRSP 195
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Wittekind M, Weinheimer S,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                2000WO-US00345
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           99US-0115271
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             WPI; 2000-465976/40.
                                                                                                                                                                                                                                                                                 195 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          N-PSDB; AAA73331
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         MO200040707-A1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               06-JAN-2000;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           08-JAN-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             19-DEC-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     13-JUL-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Synthetic
                                                                                                                                                                                                                                                                                   Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                          61
                                                                                                                                                                                                                                                                                                                                                                                                                                                     61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  AAB15222;
                                                                                                                                                                                                                                                                                                              Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                21
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                181
                                                                                                                                                                                                                                                                                                                            Best Loca
Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     a
                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Š
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          8
```

```
The present sequence is a mutated version of a fusion protein created using the Hepatitis C virus (HVV) NS3 and NS4A protease enzymes. These proteins are both essential for the replication of the virus, acting to cleave its replicative proteins from the polyprotein produced from the HCV genome. Inhibitors of the two proteins should be effective as antiviral treatments of the two proteins should be effective as antiviral treatments of the vinfection. This is useful as HCV can lead to chronic liver disease such as cirthosis, liver failure and liver cancer. The present invention concerns a number of NS3 mutants and NS3-NS4A well as enabling structural studies of the protease and proteins which can be used to identify inhibitors of this type, as well as enabling structural studies of the protease and
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            119 LVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVCTRGVAK 178
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              59 TCINGVCWTVYHGAGTRTIASPKGPVIQMYTNVDKDLVGWPAPQGSRSLTPCTCGSSDLY 118
                                                                                                                                                                                                                                                                                                                                                                                                                                                                          9
             Modified hepatitis C virus (HCV) NS3 protease comprising at least 1 substitution of a hydrophobic alpha-helix 0 amino acid to a hydrophilic amino acid, useful for screening inhibitors that may treat hepatitis C
                                                                                                                                                                                                                                                                                                                                                                                                                                                              1 MKKKGSVVIVGRINLSGDTAYAQQTRGEGCQETSQTGRDKNQVEGEVQIVSTATQTFLA
                                                                                                                                                                                                                                                                                                                                                                                                                                           1 MKKKGSVVIVGRIVLNG - - AYAQQTRGEEGCQETSQTGRDKNQVFGEVQIVSTAAQTFLA
                                                                                                                                                                                                                                                                                                                                                                                                             Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Hepatitis; NS3 protease; viral replication; chronic liver disease;
liver failure; liver cancer; mutant; mutein.
                                                                                                                                                                                                                                                                                                                                                                                                           ~
                                                                                                                                                                                                                                                                                                                                                                          Length 197;
                                                                                                                                                                                                                                                                                                                                                                                                           4; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Goldfarb V;
                                                                                                                                                                                                                                                                                                                                                                          Score 981; DB 21;
Pred. No. 4.1e-94;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Hepatitis C virus NS4A-NS3 fusion protease #8
                                                                                                                                                                                                                                                                                                                                                                                                          1; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Zhang Y,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Ž
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                8
                                                                                             Claim 23; Fig 14; 66pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            AAB15226 standard; protein; 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           179 AVDFIPVESLETTMRSP 195
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          AVDFIPVESLETIMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                (BRIM ) BRISTOL-MYERS SQUIBB
                                                                                                                                                                                                                                                                                                                                                                            96.1%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ŝ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              06-JAN-2000; 2000WO-US00345
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               99US-0115271
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Weinheimer
                                                                                                                                                                                                                                                                                                                                                                                                           Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             WPI; 2000-465976/40.
                                                                                                                                                                                                                                                                                                                                                                         Query Match
Best Local Similarity
Matches 190; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Hepatitis C virus
                                                                                                                                                                                                                                                                                                                                             197 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            WO200040707-A1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Wittekind M,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               08-JAN-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           19-DEC-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             13-JUL-2000
                                                                                                                                                                                                                                                                                                                                             Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Synthetic
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             AAB15226;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             RESULT 5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              AAB15226
                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
```

us-09-965-594-12.rag

```
2000-465976/40.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           RESULT 7
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         AAB15224
g
                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                             ŏ
                                                                                                                     The present sequence is a mutated version of a fusion protein created using the Hepatitis C virus (HCV) NS3 and NS4A protease enzymes. These proteins are both essential for the replication of the virus, acting to cleave its replicative proteins from the polyprotein produced from the HCV genome. Inhibitors of the two proteins should be effective as antiviral treatments of HCV infection. This is useful as HCV can lead to chronic liver disease such as cirrhosis, liver failure and liver cancer. The present invention concerns a number of NS3 mutants and NS3-NS4A fusion proteins which can be used to identify inhibitors of this type, as well as enabling structural studies of the protease and protease: This sequence contains the alpha-helix0 wild-type sequence.
                                                                                                                                                                                                                                                                                                                                                                                                                                              118
                                                                                                                                                                                                                                                                                                                                                                                                                                                             61 TCINGVCWTVYHGAGTRTIASPKGPVIQMYINVDKDLVGWPAPQGSRSLTPCTCGSSDLY 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   121 LVTRHADVIPVRRRGDSRGSILSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVCTRGVAK 180
                                                                                                                                                                                                                                                                                                                                                                                        28
                                                                                                                                                                                                                                                                                                                                                                                                                  9
                          Modified hepatitis C virus (HCV) NS3 protease comprising at least 1 substitution of a hydrophobic alpha-helix 0 amino acid to a hydrophilic amino acid, useful for screening inhibitors that may treat hepatitis C
                                                                                                                                                                                                                                                                                                                                                                                        1 MKKKGSVVIVGRIVLNG--AYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVSTAAQTFILA
                                                                                                                                                                                                                                                                                                                                                                                                          TCINGVCWTVYHGAGTRTIASPKGPVIQMYTNVDKDLVGWPAPQGSRSLTPCTCGSSDLY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     LVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVCTRGVAK
                                                                                                                                                                                                                                                                                                                                                               Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Hepatitis; NS3 protease; viral replication; chronic liver disease;
liver failure; liver cancer; mutant; mutein.
                                                                                                                                                                                                                                                                                                                                                             5:
                                                                                                                                                                                                                                                                                                                                  Length 197;
                                                                                                                                                                                                                                                                                                                                                             Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Goldfarb V;
                                                                                                                                                                                                                                                                                                                               Score 959; DB 21;
Pred. No. 8.2e-92;
1; Mismatches 6;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Hepatitis C virus NS4A-NS3 fusion protease #5.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Zhang Y,
                                                                                                 Example 5; Fig 18; 66pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         AAB15223 standard: protein; 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           AVDFIPVESLETTMRSP 195
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          AVDFIPVESLETIMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  (BRIM ) BRISTOL-MYERS SQUIBB
                                                                                                                                                                                                                                                                                                                                  93.98;
95.48;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Weinheimer S,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            06-JAN-2000; 2000WO-US00345
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             19-DEC-2000 (first entry)
                                                                                                                                                                                                                                                                                                                                               Best Local Similarity 95.4 Matches 188; Conservative
                                                                                                                                                                                                                                                                                                         197 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Hepatitis C virus
Synthetic.
N-PSDB; AAA73335
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Hepatitis; NS3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      WO200040707-A1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Wittekind M,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      08-JAN-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                13-JUL-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    AAB15223;
                                                                                                                                                                                                                                                                                                                                                                                                                                              59
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   119
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          179
                                                                                                                                                                                                                                                                                                          Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    181
                                                                                                                                                                                                                                                                                                                                   Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             AAB15223
AAB15223
AAB15223
AAB15223
AAB15223
AAB1523
                                                                                                                                                                                                                                                                                                                                                                                                                 a
ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Ω
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
```

```
The present sequence is a mutated version of a fusion protein created using the Hepatitis C virus (HCV) NG3 and NS4A protease enzymes. These proteins are both essential for the replication of the virus, acting to cleave its replicative proteins from the polyprotein produced from the HCV genome. Inhibitors of the two proteins should be effective as antiviral treatments of HCV infection. This is useful as HCV can lead to chronic liver disease such as cirrhosis, liver failure and liver cancer. The present invention concerns a number of NS3 mutants and NS3-NS4A fusion proteins which can be used to identify inhibitors of this type, as well as enabling structural studies of the protease and protease:inhibitor complexes. This sequence contains the alpha-helix0-liveriant.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              118
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               119 LVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVCTRGVAK 178
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           28
                                                                             Modified hepatitis C virus (HCV) NS3 protease comprising at least 1 substitution of a hydrophobic alpha-helix 0 amino acid to a hydrophilic amino acid, useful for screening inhibitors that may treat hepatitis C
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            TCINGVCWTVYHGAGTRTIASPKGPVIQMYTNVDKDLVGWPAPQGSRSLTPCTCGSSDLY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1 MKKKGSVVIVGRIVLNG--AYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVSTAAQTFLA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Hepatitis; NS3 protease; viral replication; chronic liver disease;
liver failure; liver cancer; mutant; mutein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    5;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Length 197;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Goldfarb V;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Score 951; DB 21;
Pred. No. 5.6e-91;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Hepatitis C virus NS4A-NS3 fusion protease #6.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Pred. No. 5.6e
1; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Zhang Y,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Ź
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (BRIM ) BRISTOL-MYERS SQUIBB CO.
                                                                                                                                                                                                                                                                                                       Claim 23; Fig 15; 66pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             AAB15224 standard; protein; 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          179 AVDFIPVESLETIMRSP 195
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            93.18;
94.98;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Weinheimer S,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               06-JAN-2000; 2000WO-US00345
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 19-DEC-2000 (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Best_Local Similarity 94.9
Matches 187; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 197 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Hepatitis C virus
N-PSDB; AAA73332.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       WO200040707-A1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Wittekind M,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          08-JAN-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             13-JUL-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Synthetic
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              59
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        AAB15224;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  THE STATE OF STATE OF
```

```
The present sequence is a mutated version of a fusion protein created using the Hepatitls C virus (HeV) NS3 and NS4A protease enzymes. These proteins are both essential for the replication of the virus, acting to cleave its replicative proteins from the polyprotein produced from the HCV genome. Inhibitors of the two proteins should be effective as antiviral treatments of the two proteins should be effective as chronic liver disease such as cirthosis, liver failure and liver cancer. The present invention concerns a number of NS3 mutants and NS3-NS4A fusion proteins which can be used to identify inhibitors of this type, as well as enabling structural studies of the protease and protease:Inhibitor complexes. This sequence contains the alpha-helix0-7
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                TCINGVCWTVYHGAGTRTLASPKGPV1QMYTNVDKDLVGWPAPQGSRSLTPCTCGSSDLY 118
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       LVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVCTRGVAK 178
                                                                                                                                                                                                                                                                                                                                                                                                                                                      n)
G)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Modified hepatitis C virus (HCV) NS3 protease comprising at least 1 substitution of a hydrophobic alpha-helix 0 amino acid to a hydrophilic amino acid, useful for screening inhibitors that may treat hepatitis \mathbb C
                                                                                                                                                                                                                                                                                                                                                                                                                                                  1 MKKKGSVVIVGRIVLNG--AYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVSTAAQTFLA
                                                                                                                                                                                                                                                                                                                                                                                                                        Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Hepatitis; NS3 protease; viral replication; chronic liver disease;
liver failure; liver cancer; mutant; mutein.
                                                                                                                                                                                                                                                                                                                                                                                                                        2;
                                                                                                                                                                                                                                                                                                                                                                                     Score 939; DB 21; Length 197;
Pred. No. 1e-89;
1; Mismatches 8; Indels 7;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Hepatitis C virus NS4A-NS3 fusion protease #7.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Æ.
                                                                                                                                  Claim 23; Fig 16; 56pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        AAB15225 standard; protein; 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           (BRIM ) BRISTOL-MYERS SQUIBB CO.
                                                                                                                                                                                                                                                                                                                                                                                                                    3,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 AVDFIPVESLETIMRSP 195
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    AVDFIPVESLETTMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                          92.0%;
93.4%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             99US-0115271.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               06-JAN-2000; 2000WO-US00345
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  19-DEC-2000 (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                      184; Conservative
               2000-465976/40
                                                                                                                                                                                                                                                                                                                                                                                          Query Match
Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Hepatitis C virus Synthetic.
                                                                                                                                                                                                                                                                                                                                                             197 AA;
                               N-PSDB; AAA73333
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     WO200040707-A1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             08-JAN-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   13-JUL-2000.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Hepatitis;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       119
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   179
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      AAB15225;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              59
                                                                                                                                                                                                                                                                                                                                                               Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               181
                                                                                                                                                                                                                                                                                                                                 variant
                                                                                                                                                                                                                                                                                                                                                                                                                        Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          RESULT 8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                             q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               đ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
```

```
The present sequence is a mutated version of a fusion protein created using the Hepatitis C virus (HCV) NG3 and NS4A protease enzymes. These proteins are both essential for the replication of the virus, acting to cleave its replicative proteins from the polyprotein produced from the HCV genome. Inhibitors of the two proteins should be effective as antiviral treatments of HCV infection. This is useful as NCV can lead to chronic liver disease such as cirrhosis, liver failure and liver cancer. The present invention concerns a number of NS3 mutants and NS3-NS4A fusion proteins which can be used to identify inhibitors of this type, as well as enabling structural studies of the protease and protease: Inhibitor complexes. This sequence contains the alpha-helix0-7
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    NS3 catalytic domain; NS4A peptide; NS4A-NS3 fusion construct; diagnosis; serine protease; trypsin family; screening; anti-viral compound; treatment; inhibitor; therapeutic.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            59 TCINGVCWTVYHGAGTRTIASPKGPVIQMYTNVDKDLVGWPAPQGSRSLTPCTCGSSDLY 118
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                9
                                                                     Modified hepatitis C virus (HCV) NS3 protease comprising at least 1 substitution of a hydrophobic alpha-helix 0 amino acid to a hydrophilic amino acid, useful for screening inhibitors that may treat hepatitis C
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1 MKKKGSVVIVGRIVLNG--AYAQOTRGEEGCQETSQTGRDKNQVEGEVQIVSTAAQTFLA
                                                                                                                                                                                                                                                                                                                                                                                                                                       Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1..12
/label= NS4A_peptide_1
/note= "Covalently attached to amino terminus of NS3
catalytic domain"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            /note= "Wild type Proline is replaced with Lysine"
                                                                                                                                                                                                                                                                                                                                                                                                                                     5;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Hepatitis C virus NS4A-NS3 catalytic domain fusion protein-1.
                                                                                                                                                                                                                                                                                                                                                                                                    Length 197;
                                                                                                                                                                                                                                                                                                                                                                                                                                     Indels
Goldfarb V;
                                                                                                                                                                                                                                                                                                                                                                                                                                 6
                                                                                                                                                                                                                                                                                                                                                                                                      Score 929; DB 21;
Pred. No. 1.1c-88;
                                                                                                                                                                                                                                                                                                                                                                                                                                    3; Mismatches
Zhang Y,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Š
                                                                                                                                               English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        AAY44728 standard; Protein; 191
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              AVDFIPVESLETTMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                   91.0%;
92.9%;
Ś
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   AVDFIPVESLETTMRSP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 04-MAY-2000 (first entry)
Weinheimer
                                                                                                                                                                                                                                                                                                                                                                                                                                  Matches 183; Conservative
                                                                                                                                              23; Fig 17; 66pp;
                            WPI; 2000-465976/40.
                                                                                                                                                                                                                                                                                                                                                                                                                      Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                           197 AA;
                                          N-PSDB; AAA73334
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Misc-difference
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            WO200001718-A2
Wittekind M,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 179
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            181
                                                                                                                                                                                                                                                                                                                                                                            Sednence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      AAY44728;
                                                                                                                                                                                                                                                                                                                                                                                                       Query Match
                                                                                                                                                                                                                                                                                                                                              variant
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Peptide
                                                                                                                                               Claim
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          AAY44728
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          RESULT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ç
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ά
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          qq
```

ø

```
52 AAQTFLATCINGVCWTVYHGAGTRTIASPKGPVIQMYTNVDKDLVGWPAPQGSRSLTPCT 111
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         112 CGSSDLYLVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAV 171
                                                                                                                                                                                                                                                                                                                                      This sequence represents the entire hepatitis C virus polyprotein. HCV is a member of the flavivirus family and appears to encode a basic polypeptide domain ("C") at the N-terminal of the viral polyprotein, followed by two glycoprotein domains ("EI", "E2/NSI"), upstream of the nonstructural genes NS2 through NS5. See also AAQ39134-48, AAR33982-4008 and AR38088-89.
(Updated on 25-MAR-2003 to correct PN field.)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      3 KKGSVVIVG---RIVLNG-----AYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVST
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     HCV genomic amino acid sequence isolated from infected human LG.
                                                                                                                                                                                                                            Immuno-reactive hepatitis C virus polypeptide compsns. - contg. at least 2 sequences from the first variable domain of distinct
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                DB 14; Length 2816;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Hepatitis C Virus; Non-A, non-B hepatitis Virus; HCV: NANBHV; human growth hormone; HGH; secretion signal; fusion protein;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              88.0%; Score 898.5; DB 14;
ilarity 85.8%; Pred. No. 6.1e-84;
Conservative 9; Mismatches 9;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   $
                                                                                                                                                                                                                                                                                                    Disclosure; Fig 9; 106pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   AAR40120 standard; Protein; 3011
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         93WO-0S00907
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             92US-0830024
                                      92WO-US07683
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (updated)
(first entry)
                                                                                                                                                    Weiner AJ;
                                                                                                                                                                                    WPI; 1993-117468/14.
                                                                                                              (CHIR ) CHIRON CORP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Best Local Similarity
Matches 175; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          2816 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Hepatitis C Virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         29-JAN-1993;
                                                                                                                                                                                                                                                                  HCV isolates
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           W09315193-A1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             31-JAN-1992;
                                    11-SEP-1992;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               25-MAR-2003
                                                                                                                                                  Houghton M,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             27 - JAN-1994
01-APR-1993
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1185
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         AAR40120;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Seguence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     172
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   vaccine
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 RESULT 11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   AAR40120
qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       οy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                    The present protein sequence is the fusion polypeptide, comprising the hepatitis C virus NSAA peptide-1 fragment, covalently attached to the amno terminus of NS3 catalytic domain. This fusion polypeptide contains the NS3 domain expressed in a stable. Soluble form. This fectlitates the use of the polypeptide in direct screening of potential anti-viral compounds, that are used for diagnosis and treatment of hepatitis C virus infection. It is also used to screen for inhibitors of serine protease activity. The polynucleotides are also useful to identify diagnostic or therapeutic compounds and for recombinant production of the fusion
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           WIVYHGAGIRTIASPKGPVIQMYINVDKDLVGWPAPQGSRSLIPCTCGSSDLYLVTRHAD 125
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   126 VIPVRRRGDSRGSIJSPRPISYLKGSSGGPLIGPAGHAVGTFRAAVCTRGVAKAVDF1PV 185
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SVVIVGRIVLS-KYAQQTRGLLGCIITSLTGRDKNQVEGEIQIVSTAAQTFLATCVNGVR 60
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  65
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             121 VIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVSIFRAAVCTRGVAKAVDFIPV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SVVIVGRIVLNGAYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVSTAAQTFLATCINGVC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Polymerase chain reaction; PCR: amplify; primer: hepatitis C virus; 
PCV: asymptomatic; chronically infected; epitope; viral isolate; 
domain; immunological; cross reactive; envelope protein; vaccine; 
qp53(BVDV)/qp55; hog cholera virus; pestivirus; NSI; flavivirus.
                                                                                                                                                                                                                                            Novel polypeptide comprising hepatitis C virus NS4A and NS3 domains, useful for screening for compounds useful for the diagnosis and treatment of hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          DB 21; Length 191;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      88.1%; Score 899.5; DB 21; 92.1%; Pred. No. 1.3e-85; Live 6; Mismatches 8;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               AAR34009 standard; Protein; 2816 AA
                                                                                                                                                                                                                                                                                                                        Claim 4; Fig 2; 30pp; English
                                                                                            98US-0091675
                                                      99WO-US15035
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Query Match
Best Local Similarity 92.13
Matches 175; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ESLETTMRSP 195
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ENLETTMRSP 190
                                                                                                                                                                    Bukhtiyarova
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           (updated)
                                                                                                                                (UYFL ) UNIV FLORIDA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 HCV-1 polyprotein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    191 AA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         W09306126-A1
                                                      02-JUL-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         25-MAR-2003
26-JUL-1993
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                polypeptide.
                13-JAN-2000
                                                                                                                                                                    Dunn BM,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           99
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             186
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               AAR34009;
```

RESULT 10 AAR34009

g

ò

q

q ö

ò g ò

```
172 CTRGVAKAVDFIPVESLETTMRSP 195
                                                                                                                                                                                                                                                                                                                                                                                                                                                          ż
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   AAP90158 standard; protein; 1786
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              87US-0122714.
87US-0139886.
88US-0161072.
                                                                        Kuo G;
88US-0161072.
88US-0191263.
88US-0263584.
88US-0271450.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           88GB-0027024
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           (updated)
(first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Hepatitis C virus; vaccine
                                                                                                                      Purified hepatitis C virus
                                                                                                                                                                                                                                                                                           Best Local Similarity 85.8
Matches 175; Conservative
                                                                       Houghton M, Choo QL,
                                                                                          WPI; 1989-159274/22.
                                                                                                                                                                                                                                                               1766 AA;
                                                                                                    N-PSDB; AAN92097
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Pan troglodytes
                                                  (CHIR ) CHIRON
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              18-NOV-1987;
30-DEC-1987;
26-FEB-1988;
26-FEB-1988;
06-MAY-1988;
26-OCT-1988;
14-NOV-1988;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           18-NOV-1988;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          25-MAR-2003
10-NOV-1989
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   GB2212511-A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        26-JUL-1989
                                                                                                                                                                                                                                                                Sedneuce
                                                                                                                                                                                                                                                                                                                                                                                                                                                                   469
                                                                                                                                                                                                                                                                                                                                                                25
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       AAP90158;
                                                                                                                                                                                                                                                                                  Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         AAP90158
                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                 ð
                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                   .;
?
                                                                                                                                                                                                                                                                                                                                                                                             Sequence encoded in the hepatitis C virus (HCV) cDNA inserts in clones 141, 11b, 7f, 7e, 8h, 33c, 40b, 37b, 35, 36, 81, 32, 33b, 25c, 14c, 8f, 33f, 33g and 39c.
                                                                                                                                                                                                                                                                                                              3 KKGSVVIVG----RIVLNG------AYAQQIRGEEGCQETSQIGRDKNQVEGEVQIVSI
                                                                                                                                     RNA was isolated from the plasma of a HCV scropositive human (designated "LG") and cDNA was prepared from it. The cDNA was PCR amplified using specific primers with sequences based on the prototype HCV-1 cDNA sequence (GENBANK M62321). Further amplification using nested primers resulted in 7 adjacent HCV DNA fragments which could be assembled into a full-length sequence. The DNA sequence was determined and translated into the genomic amino acid sequence. Comperison of the LG genomic amino acid sequence with that from HCV-1 showed 134 amino acid differences. (Updated on 25-MAR-2003 to correct PN field.)
                                                                                                                                                                                                                                                                      87.7%: Score 895.5; DB 14; Length 3011; 85.8%: Pred. No. 1.4e-83; Live 8; Mismatches 10; Indels 11;
                                                                                plasmid pHCV-162 is a mammalian expression systems for HCVI eeins - useful for diagnosing HCV infection and as vaccines
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Hepatitis C virus (HCV); non-A, non-B hepatitis (HANBH)
                              Frail
                              Devare SG,
                                                                                                                                                                                                                                                                                                                                                                                                                                              Example 1; Page 39-49; 100pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         AAP92041 standard; protein; 1766 AA
                              Desai SM,
                                                                                                  for preventing HCV infection
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     88EP-0310922
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        87US-0122714.
87US-0139886.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                25-MAR-2003 (updated)
02-MAR-1990 (first entry)
                                                                                                                                                                                                                                                                                 Best Local Similarity 85.8°
Matches 175; Conservative
                                        Zeck BJ;
                                                           WPI; 1993-258673/32.
                              Casey JM,
                                                                                                                                                                                                                                                     3011 AA;
          (ABBO ) ABBOTT LAB
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Repatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        18-NOV-1987;
30-DEC-1987;
                                        Yamaguchi J,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 31-MAY-1989
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            EP318216-A
                                                                                                                                                                                                                                                      Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             AAP92041;
                                                                                                                                                                                                                                                                        Query Match
                              SI,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
 ò
                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                             õ
                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                     õ
```

```
It is the sequence encoded in the open reading frame of hepatitis C virus CoDA Inserts in clones 141,m 11b, 7f, 7e, 8h, 32c, 40b, 37b, 35, 36, 81, 32, 33b, 25c, 14c, 8f, a31f, 33g and 39c. It is antigenic and could be used in immunoassay reagents and vaccines and to generate antibodies useful in diagnosis and passive immunotherapy for HCV infection/non-A, (Updated on 25-WAR-2003 to correct PR field.)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       289 RRGREILLGPADGMVSKGWRLLAPITAYAQQTRGLLGCIITSLTGRDKNQVEGEVQIVST 348
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         111
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        171
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              AAQTFLATCINGVCWTVYHGAGTRTIASPKGPVIQMYTNVDKDLVGWPAPQGSRSLTPCT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  112 CGSSDLYLVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   3 KKGSVVIVG---RIVLNG-----AYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVST
                                                                  Claim 13; Figure 26-1, 26-2, 26-3, 26-4, 26-5, 26-6; 139pp; English
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 DB 10; Length 1766;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Protein sequence of hepatitis c virus composite cDNA
and associated nucleic acids and polypeptide(s)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              87.6%; Score 894.5; DB 10;
85.8%; Pred. No. 8.6e-84;
tive 8; Mismatches 10;
```

œ

2

11;

Indels

Length 2261;

XXX XXX BET X BET

```
380 RRGREILLGPADGMVSKGWRLLAPITAYAQQTRGLIGCIITSLTGRDKNQVEGEVQIVST 439
                                                                                                                                                                                                                                                                                                                                                                                                                                                                             112 CGSSDLYLVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAV 171
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Sequence encoded in the hepatitis C virus (HCV) cDNA inserts in clones 12f\ \text{through}\ 15e.
                                                                                                                                                                                                                                                                                                                                                                                                                                 AAQTFLATCINGVCWTVYHGAGIRTIASPKGPVIQMYTNVDKDLVGWPAPQGSRSLTPCT
                                                                                                                                                                                                                                                                                                                                                  3 KKGSVVIVG---RIVLNG----AYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVST
                                                                                                                                                     The sequence is the peptide encoded by the composite hepatitis C virus (HCV) cDNA of AAN90331. The polypeptides are used to diagnose HCV-induced NANH+, to raise antibodies for immunoassay or treatment, our to produce vaccines.
(Updated on 25-MAR-2003 to correct PR field.)
                                                         Hepatitis C virus gene - used for prodn. of polynucleotide polypeptide(s) and antibodies for diagnosis, prevention and treatment of infection.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Hepatitis C virus (HCV); non-A, non-B hepatitis (HANBH)
                                                                                                                                                                                                                                                                                     DB 10;
                                                                                                                                                                                                                                                                                  Ouery Match 87.6%; Score 894.5; DB 10; Best Local Similarity 85.8%; Pred. No. 1.2e-83; Matches 175; Conservative 8; Mismatches 10;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               172 CTRGVAKAVDFIPVESLETIMRSP 195
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      AAP92047 standard; protein; 2301 AA
                                                                                                                          Disclosure; fig 32; 235pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Kuo G;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       87US-0122714.
87US-0139886.
88US-0161072.
88US-0191263.
88US-0263584.
88US-0271450.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            88EP-0310922
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (updated)
(first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Roughton M, Choo QL,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           (CHIR ) CHIRON CORP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         WPI; 1989-159274/22
             WPI; 1989-215054/30
                                                                                                                                                                                                                                                     2261 AA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Hepatitis C virus
                            N-PSDB; AAN90331
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         18-NOV-1987;
30-DEC-1987;
26-FEB-1988;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        06-MAY-1988;
26-OCT-1988;
14-NOV-1988;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            18-NOV-1588;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     25-MAR-2003
02-MAR-1990
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              31-MAY-1989
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              EP318216-A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      AAP92047;
                                                                                                                                                                                                                                                     Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       AAP92047
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       6
                                                                                                                                                                                                                                                                                                                                                                                                                                             q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       a
 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                             δ
                                                                                                                                                                                                                                                                                                                                                                                               3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        408
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             52 AAQTFLATCINGVCWTVYHGAGTRTIASPKGPVIQMYTNVDKDLVGWPAPQGSRSLTPCT 111
                                                                                                                                                                                                                                   The sequence is encoded by the composite CDNA of AAN90327. These antigens react with antibodies in patients with non-A non-B hepatitis (NANBH). They can be used to diagnose HVV-induced NANBH, to raise antibodies for immunoassay or treatment, or to produce vaccines. (Updated on 25-MAR-2003 to correct PR field.)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Hepatitis C virus gene - used for prodn. of polynucleotide probes, polypeptide(s) and antibodies for diagnosis, prevention and treatment of infection.
                                                                                                                                                                                                                                                                                                                                                                                                                             3 KKGSVVIVG---RIVLNG-----AYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVST
                                                                                                                                                                                                                                                                                                                                                                                               11; Gaps
                                                                                                                                                                                                                                                                                                                                                             Query Match 87.6%; Score 894.5; DB 10; Length 1786; Best Local Similarity 85.8%; Pred. No. 8.7e-84; Matches 175; Conservative 8; Mismatches 10; Indels 11;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Hepatitis C virus; clone 12f; clone 15e; probe; vaccine.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Peptide encoded by composite hepatitis C virus cona.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CIRGVAKAVDFIPVESLETTMRSP 195
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         AAP90164 standard; protein; 2261 AA
                                                                                                                                                                                                        Disclosure; fig 26; 30pp; English
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     87US-0122714.
87US-0139886.
88US-0161072.
88US-0263584.
                                                            ::
::
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         88GB-0027024
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        (CHIR ) CHIRON CORPORATION.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Choo QL, Kuo
88US-0263584
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                (updated)
(first entry)
                              (CHIR ) CHIRON CORPORATION
                                                            Choo OL,
                                                                                         WPI; 1989-215054/30,
N-PSDB; AAN90327.
                                                                                                                                                                                                                                                                                                                                   Sequence 1786 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Pan troglodytes
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        30-DEC-1987;
26-FEB-1988;
26-OCT-1988;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         18-NOV-1988;
26-OCT-1988;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         18-NOV-1987;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               25-MAR-2003
01-NOV-1989
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          GB2212511-A
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Houghton M,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         26-JUL-1989
                                                          Houghton M,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  AAP90164;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        172
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     469
```

QQ

g ò q

g õ

õ

N-PSDB; AAN92103

```
It is the sequence encoded in the open reading frame of hepatitis C virus (HCV) cbMx inserts in clones 12f through 15c. It is antiqenic and could be used in immunoassay reagents and vaccines and to generate antibodies useful in diagnosis and passive immunotherapy for HCV infection/non-A, non-B hepatitis.

(Updated on 25-WAR-2003 to correct PR field.)
                                                                                                                                                                                                                                                                                                                                                                  3 KKGSVVIVG---RIVING-----AYAQQTRGEEGCQETSOTGRDKNOVEGEVQTVST 51
                                                                                                                                                                                                                                                                               11; Gaps
                                                                                                                                                                                                                                                Query Match 87.6%; Score 894.5; DB 10; Length 2301; Best Local Similarity 85.8%; Pred. No. 1.2e-83; Matches 175; Conservative 8; Mismatches 10; Indels 11;
                             Purified hepatitis C virus - and associated nucleic acids and polypeptide(s) % \left( \frac{1}{2}\right) =0
                                                                          Claim 13; Figure 32-1 - 32-7; 139 pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CTRGVAKAVDFIPVESLETIMRSP 195
                                                                                                                                                                                                                      Sequence 2301 AA;
q
                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                      qq
                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      90
```

Search completed: August 30, 2003, 19:12:22 Job time: 47.1697 secs

CIRGVAKAVDFIPVENLETIMRSP 583

```
GenCore version 5.1.6 Copyright (c) 1993 - 2003 Compugen Ltd.
```

OM protein - protein search, using sw model

Run on:

August 30, 2003, 19:02:22; Search time 16.0488 Seconds (without alignments) 1168.492 Million cell updates/sec

Title: Perfect score:

US-09-965-594-12 1021 1 MKKKGSVVIVGRIVLNGAYA.....VAKAVDFIPVESLETIMRSP 195 Sequence:

Scoring table: BLOSUM62 Gapop 10.0 , Gapext 0.5

283308 seqs, 96168682 residues Searched:

Total number of hits satisfying chosen parameters:

283308

Minimum DB seq length: 0 Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0% Maximum Match 100% Listing first 45 summaries

PIR\_76:\*
1: pir1:\*
2: pir2:\*
3: pir3:\*
4: pir4:\* Database :

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Description		qenome polyprotein												polyprotein - dour	hypothetical prote	_	ā	nitrate/nitrite se	nitrate/nitrite se	env polyprotein -	insulin-like growt	hypothetical prote	envelope surface q	aromatic	beta transducin ho	probable periplasm	centromere protein	hypothetical prote		-
SUMMARIES		GNWVC3	S40770	GNWVCH	GNWVIW	GNWVTC	GNWVCJ	A45573	S18030	JC5620	JQ1303	GNWVJB	T08839	T08841	B71360	AB1775	139383	B81104	C81911	VCMVFG	IOHUI	T49806	A46165	H83144	T18234	B71284	S28261	T34513	T06118	E95261
DB	-	Н		-	-	٦	_		_	-	_	٦	7	7	7	~	~	7	7		7	~	~	~	7	7	_	7	7	7
Length		3011	3011	3011	3010	3010	3010	3010	3010	3014	3033	3033	2970	3005	495	1334								209	981	398	2663	S	270	393
% Query Match		87.6	87.5	86.1	83.0	82.5	82.1	82.0	80.4	74.5	65.7	65.5	25.5	25.3	8.4	8.4	0 8	7.9	7.9	7.8	7.7	7.7	7.7	7.7	7.7	7.6	7.6	7.6	7.5	7.5
Score		894.5	893.5	879.5	847	842	838	837	821	761	670.5	668.5	260	258	85.5	85.5	$\alpha$	80.5			19	79	^	78.5			77.5			
Result No.		1	7	٣	*	S	9	7	<b>6</b> 0	6	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	56	27	28	53

serine proteinase hypothetical prote	hypothetical prote	hypothetical prote	env polyprotein pr orsellinic acid sv	zona pellucida gly	probable exported	hypothetical prote	ice nucleation pro	serine proteinase	conserved hypothet	env polyprotein .	hypothetical prote	phosphate ABC tran
B98127 C72705	A84782 T35785	S41034	VCMVSA T30871	S70401	AH1030	S23441	T18042	99668н	G87265	VCFVER	S76618	H84203
010	7	101	<b>⊣</b> ~	~	~	~	~	~	~	~	~	7
397	854	887	1293	415	492	755	1176	239	239	603	317	266
7.5	4.7	4.6	7.3	7.3	7.3	7.3	7.3	7.2	7.2	7.2	7.2	7.2
76.5	75.5	75.5	c 27	74.5	74.5	74.5	74.5	74	74	74	73.5	73.5
30 31	3.2	34	36 36	37	38	36	40	41	42	43	44	4.5

## ALIGNMENTS

A: Title: Analysis of a new hepatitis C virus type and its phylogenetic relationship A: Reference number: PQ0393; MUID:92268871; PMID:1316939 A: Recession: PQ0403 A: Molecule type: genomic RNA A: Residues: 1577-1633 < CHA> A: Cross references: DBU:010128 A: Residues: 1577-1633 < CHA> A: Cross references: DBU:010128 A: Residues: 1577-1633 < CHA> A: Cross references: DBU:010128 A: Reperimental source: isolates E-b16 A: A: Experimental source: solates E-b17 A: Residues: 1577-1633 < CH2> A: Residues: 1577-1633 < CH	Query Match 87.6%; Score 894.5; DB 1; Length 3011; Best Local Similarity 85.8%; Pred. No. 1.7e-74; Matches 175; Conservative 8; Mismatches 10; Indels 11; Gaps 2;
--	---

ò

us-09-965-594-12.rpr

```
R:Chen, P.J.; Lin, M.H.; Tal, K.F.; Liu, P.C.; Lin, C.J.; Chen, D.S.
Virology 188, 102-113, 1992
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            C; Accession: A40244
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  112
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              A:Cross-references: GB:D00831; NID:g221511: PIDN:BAA00705.1; PID:g221512

A:Experimental source: isolate HC-J1

A:Experimental source: isolate HC-J1

A:Experimental source: isolate HC-J1

C:Superfamally: hepatitis C virus genome polyprotein

C:Superfamally: hepatitis C virus predicted cRPA>

F:116-191/Product: capsid protein M #status predicted cREM>

F:116-191/Product: mayor envelope protein E #status predicted cNS1>

F:130-1204/Product: nonstructural protein NS2 #status predicted cNS2>

F:130-1206/Froduct: nonstructural protein NS2 #status predicted cNS2>

F:1316-1317/Region: nucleotide-binding motif A (P-loop)

F:1317-1317/Region: nucleotide-binding motif A (P-loop)

F:1316-1319/Region: nucleotide-binding motif B

F:1616-1862/Product: nonstructural protein NS4 #status predicted cN4A>

F:1863-2013/Product: nonstructural protein NS5 #status predicted cN4S>

F:2014-3011/Product: nonstructural protein NS5 #status predicted cN4S>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             A;COSS-references: EMBL:D10749; NID:9221586; PIDN:BAA01582.1; PID:9221587
A;COSS-references: EMBL:D10749; NID:9221586; PIDN:BAA01582.1; PID:9221587
B;OXAMOLO, H.; OXAGA, S.; Sugiyama, Y.; Yotsumoto, S.; Tanaka, T.; Yoshizawa, H.; Tsuda, Jpn. J. Exp. Med. 60, 167-177, 1990
A;Title: The 5'-terminal sequence of the hepatitis C virus genome.
A;Reference number: PC1284; MUID:91013116; PMID:2170712
A;Recession: PC1288
A;Molecule type: genomic RNA
A;Residues: 1513 <OX2>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   genome polyprotein – hepatitis C virus
N:Contains: capsid protein C; envelope protein M; hepacivirin (EC 3.4.21.98) (nonstructu
protein NS4a; nonstructural protein NS4b; nonstructural protein NS5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1005 RRGREILLGPADGMVSKGWRLLAPITAYAQQTRGLLGCIITSUTGRDKNQVEGHVQIVST 1064
                                                                                                                      1065 AAQTFLATCINGVCWTVYHGAGTRTIASPKGPVIOMYTNVDQDLVGWPAPQGSRSLTPCT 1124
                                                                                                                                                                                                                   112 CGSSDLYLVTRHADVIPVRRRCDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAV 171
                                                            52 AAQTFLATCINGVCWTVYHGAGTRIIASPKGPVIQMYTNVDKDLVGWPAPQGSRSLTPCT 111
                                                                                                                                                                                   112 CGSSDLYLVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAV 177
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             C.Species: hepatitis C virus
C.Date: 19-May-2000 *sequence_revision 19-May-2000 *text_change 19-Jan-2001
C.Accession: S40770: PC1285
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            52 AAQTFLATCINGVCWTVYHGAGTRTIASPKGPVIQMYINVDKDLVGWPAPQGSRSLTPCT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     3 KKGSVVIVG---RIVLNG----AYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVST
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           87.5%; Score 893.5; DB 1;
85.8%; Pred. No. 2.1e-74;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     7; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                R;Okamoto, H.
submitted to the EMBL Data Library, March 1992
A;Reference number: $40770
A;Accession: $40770
A;Accession: $40770
A;Accelle type: genomic RNA
A;Residues: 1-3011 <0KA>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1185 CTRGVAKAVDFIPVESLETTMRSP 1208
                                                                                                                                                                                                                                                                                                                                          CTRGVAKAVDFIPVESLETTMRSP 195
                                                                                                                                                                                                                                                                                                          CTRGVAKAVDFIPVESLETTMRSP 195
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Similarity 85.8
75; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     175;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Query Match
Best Local (
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                            a
                                                                                                                                                                                                                                   q
                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                          q
                                                         ò
                                                                                                                                                                                ò
```

```
A; Molecule type: genomic RNA
A; Residues: 1-3011 < INC.
A; Cross-references: GB:WG/4513; NID:g329737; PIDN:AAA45534.1; PID:g329738
B; Inchauspe, G:, Zebedee, S.; Lee, D.H.; Sugitani, M.; Masoff, M.; Prince, A.M.
Proc. Natl. Acad. Sci. U.S.A. 88, 10992-10296, 1991
A; Tille: Genomic structure of the human prototype strain H of hepatitis C virus: comp
A; Reference number: A41546; WID:92052256; PMID:1658800
A; Contents: annotation acid nor nucleotide sequence is given
A; Note: neither amino acid nor nucleotide sequence is given
C; Superfamily: hepatitis C virus genome polyprotein
C; Superfamily: hepatitis C virus genome polyprotein
C; Superfamily: hepatitis C virus genome polyprotein
C; Keywords: ATP; capsid protein; envelope protein; glycoprotein; hydrolase; nonstruct
F; 116-191/Product: envelope protein B *status predicted <NSI>F; 1307-139/Product: nonstructural protein NIS *status predicted <NSI>F; 1107-1615/Product: nonstructural protein NIS *status predicted <NSI>F; 1107-1615/Product: nonstructural protein NIS *status predicted <NSI>F; 11107-1117/Region: nucleotide-binding motif B
F; 1111-1117/Region: DEXH motif
F; 1111-1111-1111/Region: DEXH motif
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              F:1616-1862/Próduct: monstructural protein NS4a #status predicted <N4A>
F:1863-2013-Product: nonstructural protein NS4b #status predicted <N4B>
F:2014-3011/Product: nonstructural protein NS5 #status predicted <NS5>
F:2014-3011/Product: nonstructural protein NS5 #status predicted <NS5>
F:196,209,234,305,325,417,423,430,448,476,532,540,556,576,623,645,1213,1255,2041,2240
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    genome polyprotein - hepatitis C virus (strain Taiwan)
N;Contains: capsid protein C; envelope protein M; hepacivirin (BC 3.4.21.98) (nonstru
protein NS4a: nonstructural protein NS4b; nonstructural protein NS5
C;Species: hepatitis C virus
A;Note: host Homo sapiens (man)
C;Date: 31-Dec-1992 *sequence_revision 31-Dec-1992 *text_change 19-Jan-2001
                                                                                                                                                                                                      C.Species: hepatitis C virus
A.Note: host Homo saplens (man)
C.Date: 31-Dec-1992 #sequence_revision 31-Dec-1992 #text_change 19-Jan-2001
C.Bate: 31-Dec-1992 #sequence 19-Jan-2001
C
qenome polyprotein - hepatitis C virus (strain H)
N/Contains: capsid protein C; envelope protein M; hepacivirin (EC 3.4.21.98) (nonstruprotein NS4a; nonstructural protein NS4a; nonstructural protein NS4b; nonstructural protein NS5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                52 AAQTFLATCINGVCWTVYHGAGTRTIASPKGPVIQMYTNVDKDLVGWPAPQGSRSLTPCT 111
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CGSSDLYLVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAV 171
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        3 KKGSVVIVG---RIVLNG-----AYAQQIRGEEGCQETSQIGRDKNQVEGEVQIVST
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Length 3011;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              DB 1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           13;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       86.1%; Score 879.5; DB 1
83.8%; Pred. No. 4.1e-73;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           9; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CIRGVAKAVDFIPVESLETIMRSP 195
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Local Similarity the les 171; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          A;Reference number: A36814
A;Accession: A36814
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Similarity
```

 $\sim$ 

```
78
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Ouery Match
Best Local
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Best Loca
Matches
             Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    δý
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                8
                                                                                                                                                                                   a
                                                                                                                                                                                                                                                                                                                                                              셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        A;Title: Structure and organization of the hepatitis C virus genome isolated from human A;Reference number: A38465, MUD:91140698; PMID:1847440
A;Recession: A38465
A;Rocession: A38465
A;R
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              F.1616-1862/Product: nonstructural protein NS4a *status predicted <N4A>
F.1863-2013/Product: nonstructural protein NS4b *status predicted <N4B>
F.2014-3010/Product: nonstructural protein NS5 *status predicted <NS5>
F.196,209,234,250,305,325,417,423,430,448,532,540,556,576,623,645,1213,1255,2041,2077,22
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          N;Contains: capsid protein C; envelope protein M; hepacivirin (EC 3.4.21.98) (nonstructut protein NS4) wonstructural protein NS5 protein NS5 c; Species: hepatitis C virus C virus C; Date: 31-Mar-1992 *sequence_revision 31-Mar-1992 *text_change 19-Jan-2001 C; Date: 31-Mar-1992 *sequence_revision 31-Mar-1992 *text_change 19-Jan-2001 C; Accession: A3465 C; Accession: A3465 C; Fujta, A: Manabe, S:; Murakami, S:; Fujita, J:; Onishi, E:; J: Virol. 65, 1105-1113, 1991
                                                                                                                                                                                                                    A:Cross-references: GB.M84754
C:Superfamily: hepatitis C virus genome polyprotein
C:Superfamily: hepatitis C virus genome polyprotein; hydrolase; nonstructura
C:Superfamily: hepatitis C virus genome protein; hydrolase; nonstructura
C:Superfamily: hepatitis C virus genome protein; hydrolase; nonstructura
E:115.Pproduct: capsid protein C *status predicted <EPM>
F:116.191/Product: envelope protein E *status predicted <MEE>
F:390-729/Product: nonstructural protein NS1 *status predicted <NS2>
F:730-1006/Product: nonstructural protein NS1 *status predicted <NS2>
F:1230-1237/Region: nucleotide-binding motif A (P-loop)
F:1317/Region: nucleotide-binding motif B
F:1316-1319/Region: DEXH motif
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      F;1863-2013/Product: nonstructural protein NS4b *status predicted <NS5>
F;2014-3010/Product: nonstructural protein NSS *status predicted <NS5>
F;196,209,233,234,250,305,325,417,423,430,448,532,540,556,576,623,645,1213,1255,2041,207
A,Title: The Taiwanese hepatitis C virus genome: sequence determination and mapping the A,Reference number: A40244; MUID:92230206; PMID:1314449
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               78 ASPKGPVIQMYTNVDKDLVGWPAPQGSRSLTPCTCGSSDLYLVTRHADVIPVRRRGDSRG 137
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    AYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVSTAAQTFLATCINGVCWTVYHGAGTRTI 77
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVCTRGVAKAVDFIPVESLETTMRSP 195
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             F;1616-1862/Product: nonstructural protein NS4a #status predicted <N4A>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 .;
O
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Length 3010;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Length 3010;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 10; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Score 842; DB 1;
Pred. No. 1.3e-69;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Score 847; DB 1;
Pred. No. 4.3e-70;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             11; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              polyprotein - hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                82,5%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          83.0%;
88.2%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Local Similarity 88.2%
nes 157; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              F;1316-1319/Reqion: DEXH motif
                                                                                                                                          A; Molecule type: genomic RNA
A; Residues: 1-3010 <CHE>
                                                                                                A; Accession: A40244
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     genome
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         à
```

Local Similarity

```
A.Experimental source: Japanese isolate
C.Comment: The cleavage sites of this polyprotein have not been determined.
C.Comment: The cleavage sites of this polyprotein
C.Superfamily: hepatitis C virus genome polyprotein
C.Superfamily: hepatitis C virus genome polyprotein;
C.Superfamily: hepatitis C virus genome polyprotein;
F.2-115/Product: capsid protein C *status predicted <CPC>
F.10-19/Product: envelope protein M *status predicted <MEE>
F.19-19/Product: movelope protein E *status predicted <MEE>
F.30-129/Product: nonstructural protein NSI *status predicted <NSI>F.30-129/Product: nonstructural protein NSI *status predicted <NSI>F.130-1306/Product: hepativirin *status predicted <NSI>F.1310-1317/Region: nucleotide-binding motif A (P-loop)
F.1315-1317/Region: DEXH motif
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   F;1616-1862/Product: nonstructural protein NS4a #status predicted <NVA>F;1863-2013-Product: nonstructural protein NS4b #status predicted <NVAP>F;2014-3010/Product: nonstructural protein NS5 #status predicted <NS5>F;2014-3010/Product: nonstructural protein NS5 #status predicted <NS5>F;196,209,234,250,305,325,417,423,430,448,532,556,576,623,645,1213,1255,2041,2077,2:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                A; Molecule type: genomic RNA
A; Residues: 1-3010 <a href="Arcession: Rhotology">A; Residues: 1-3010 <a href="Arcession: Rhotology">A; Rhotology (Rhotology (Rhotolog
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ö
     ó
                                                                                                                                                                                                                                                                                                            78 ASPKGPVIQMYINVDKDLVGWPAPQGSRSLIPCTCGSSDLYLVIRHADVIPVRRRGDSRG 137
                                                                                                           77
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVCTRGVAKAVDFIPVESLETTMRSP 195
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  138 SLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVCTRGVAKAVDFIPVESLETIMRSP 195
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             18 AYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVSTAAQTFLATCINGVCWTVYHGAGTRTI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ASPKGPVIQMYINVDKDLVGWPAPQGSRSLIPCTCGSSDLYLVTRHADVIPVRRRGDSRG
                                                                                                           18 AYAQQIRGEEGCQETSQIGRDKNQVEGEVQIVSIAAQIFLAICINGVCWIVYHGAGIRII
     Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Gaps
     ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            82.1%; Score 838; DB 1; Length 3010;
85.4%; Pred. No. 3e-69;
tive 16; Mismatches 10; Indels
10; Indels
     Mismatches
     14;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Conservative
154; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  A; Molecule type: genomic RNA A; Residues: 2650-2707 <KA2>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       al Similarity
152; Conserv
```

us-09-965-594-12.rpr

```
genome polyprotein - hepatitis C virus (isolate EUH1480)

W.Contains: capsid protein C; envelope protein M; hepacivirin (EC 3.4.21.9B) (nonstru protein NS44: nonstructural protein NS5
C; Species: hepatitis C virus
C; Species: hepatitis C virus
C; Species: hepatitis C virus
C; Date: 19-May-2000 #sequence_revision 19-May-2000 #text_change 19-Jan-2001
C; Accession: JC5620
R; Chamberlain, R; W; Adams, N.J.; Taylor, L.A.; Simmonds, P.; Elliott, R.M.
Biochem. Biophys. Res. Commun. 236, 44-49, 1997
A; Title: The complete coding sequence of hepatitis C virus genotype 5a, the predomina A; Reference number: JC5620; MUID:9736593; PMID:923423
A; Accession: JC5620
A; Molecule type: mRNA
A; Residues: 1-3014 CGHA>
A; Cross-references: GB:Y13184
A; Residues: 1-3014 CGHA>
A; Cross-references: GB:Y13184
A; Residues: 1-3014 CGHA>
A; Cross-references: GB:Y13184
A; Cross-references: GB:Y13184
A; Reperimental source: genome polyprotein
C; Superfamily: hepatitis C virus genome protein E *status predicted < CEPW>F:116-191/Product: envelope protein E *status predicted < MEE>F:384-400*Region: hypervariable * status predicted < NEE>F:384-400*Region: hypervariable * status predicted < NEE>F:384-400*Region: hypervariable * Status predicted < NEE>F:384-400*Region: hypervariable * Status predicted < NEE>F:386-A-70*Region: hypervariable * Sta
                                                  'V',623-624,'S',626-652,'DL',655-761,'T',763-782 <HOW>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            78 ASPKGPVIQMYINVDKDLVGWPAPQGSRSLIPCICGSSDLYLVIRHADVIPVRRRGDSRG 137
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               18 AYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVSTAAQTFLATCINGVCWTVYHGAGTRTI 77
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   F;1617-1863/Product: nonstructural protein NS4a #status predicted <N4A>F;1864-2014/Product: nonstructural protein NS4b #status predicted <N4B>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          F.390-730/Product: nonstructural protein NSI #status predicted <NSI>F.731-1007/Product: nonstructural protein NSI #status predicted <NS2>F.731-1007/Product: nonstructural protein NSI #status predicted <NS2>F.1231-1238/Region: nucleotide-binding motif A (P-loop)
F.3131-3138/Region: nucleotide-binding motif B
F.317-1320/Region: DEXH motif
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ;
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Length 3010;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     13; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Score 821; DB 1;
Pred. No. 1.1e-67;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                13; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     80.4%;
85.4%;
A;Molecule type: genomic RNA
A;Residues: 1-547,'T',549-621,'V
A;Cross-references: EMBL:X61591
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Matches 152; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Query Match
Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ô
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             A. Molecule type: DNA
A. Residues: 1-3010 cTANA
A. Residues: 1-3010 cT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      A.Variety: isolate JKI
C;Date: 19-May-2000 #text_change 23-Mar-2001
C;Date: 19-May-2000 #sequence_revision 15-May-2000 #text_change 23-Mar-2001
C;Accession: 518030; 533570; A48332; 518029
R;Honda, M.; Kaneko, S.; Masashi, U.; Kobayashi, K.; Murakami, S.
Submitted to the EMBL Data Library, September 1991
A;Description: A Whole genome of hepatitis C virus cDNA was isolated from a single patie
A;Reference number: $18028
A;Accession: $18030
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      genome polyprotein - hepatitis C virus (isolate JK1)
N'Contains: appsid protein C; envelope protein M; hepacivirin (EC 3.4.21.98) (nonstructu
protein NS4a; nonstructural protein NS4b; nonstructural protein NS5
C;Species: hepatitis C virus
                         genome polyprotein - hepatitis C virus (strain JT)
N:Contains: capsid protein C; envelope protein M; hepacivirin (EC 3.4.21.98) (nonstructuroturoturotus Notes, nonstructural protein NS5
C;Species: hepatitis C virus
C;Date: 19-May.2000 *sequence_revision 19-May-2000 *text_change 19-Jan-2001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               A.Cross-references: EMBL:X61596; NID:959478; PIDN:CAA43793.1; PID:959479
A.Experimental source: isolate JK1 from an individual
R:Honda, M.; Raneko, S.; Unoura, M.; Kobayashi, K.; Murakami, S.
Arch. Virol. 128, 163-169, 1993
A.Title: Sequence analysis of putative structural regions of hepatitis C virus isolated A;Reference number: A48332; MUID:93119270; PMID:8380322
                                                                                                                                                                                                                                                                                                                            R:Tanaka, T.; Kato, N.; Nakagawa, M.; Ootsuyama, Y.; Cho, M.J.; Nakazawa, T.; Hijikata,
Virus Res. 23, 39-53, 1992
A:Title: Molecular cloning of hepatitis C virus genome from a single Japanese carrier:
A:Reference number: A45573; MUID:92295714; PMID:1318627
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            78 ASPKGPVIQMYTNVDKDI,VGWPAPQGSRSI,TPCTCGSSD1,YLVTRHADVIPVRRRGDSRG 137
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     18 AYAQQIRGEEGCQETSQIGRDKNQVEGEVQIVSTAAQIFLAICINGVCWIVYHGAGIRII 77
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      138 SLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVCTRGVAKAVDFIPVESLETTMRSP 195
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            F;1616-1862/Product: nonstructural protein NS4a *status predicted <N4A>F;1863-2013/Product: nonstructural protein NS4b *status predicted <N4B>F;2014-3010/Product: nonstructural protein NS5 *status predicted <NNS5>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Length 3010;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          11; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Match 82.0%; Score 837; DB 1; Local Similarity 86.5%; Pred. No. 3.7e-69; es 154; Conservative 13; Mismatches 11.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          A; Molecule type: genomic RNA A; Residues: 1-3010 <HON>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         A; Status: preliminary
                                                                                                                                                                                                                                                                                    C; Accession: A45573
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  A; Accession: A45573
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    à
```

ö

Gaps

a ò a õ g

```
Genome polyprotein - hepatitis C virus (strain HC-J8)
N;Contains: capsid protein C; envelope protein M; hepacivirin (EC 3.4.21.98) (nonstranction in NS4s in nonstructural protein NS4s)
N;Contains: capsid protein C; envelope protein M; hepacivirin (EC 3.4.21.98) (nonstrains: capsid protein NS4s)
C;Species: hepatitis C virus
C;Date: 31-Dec-1992 *sequence_revision 31-Dec-1992 *text_change 19-Jan-2001
C;Accession: A40250; PQ0397; PQ0559
N;Okamoto, H.; Kurai, K.; Okada, S.I.; Yamamoto, K.; Lizuka, H.; Tanaka, T.; Fukuda, Virology 188, 331-341, 1992
A;Title: Full-length sequence of a hepatitis C virus genome having poor homology to A;Reference number: A40250; MUID:9223032; PMID:1314459
A;Reference number: A40250
A;Molecule type: genomic RNA
A;Residues: 1-3033 < CACAA
A;Residues: 1-3033 < CACAA
A;Cross-references: GB:D10988; GB:D01221; NID:g221608; PIDN:BAA01761.1; PID:g221609
R;Chan, S.W.; McOmish, F.; Holmes, E.C.; Dow, B.; Peutherer, J.F.; Follett, E.; Yap, J. Chen. Virol. 73, 1131-1141, 1992
A;Title: Analysis of a new hepatitis C virus type and its phylogenetic relationship A;Reference number: PQ0393; MUID:92268871; PMID:1316939
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        A: Molecule type: genomic RNA
A: Residues: 2678-2754 CGRA>
A: Cross-references: DBM:DB0130134
A: Experimental source: isolate E-bl2
R: Kato, N.; Ootsuyama, Y.; Ohkoshi, S.; Nakazawa, T.; Mori, S.; Hijikata, M.; Shimot
Biochem: Biophys. Res. Commun. 181, 279-285, 1991
A; Title: Distribution of plural HCV types in Japan.
A; Reference number: PQ0554; MUID:92068204; PMID:1720309
A; Accession: PQ0559
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ACCOSS-TEFEGREE (BE.101662; GB.1090518; NID:9221523; PIDN:BAA01418.1; PID:9221524
CS.Superfamily: hepatitis C virus genome polyprotein
CKEYWORDS: ATP; capsid protein; envelope protein; glycoprotein; hydrolase; nonstruc
CKEYMORDS: ATP; capsid protein C *status predicted <CPC>
F:1-115/Product: capsid protein C *status predicted <CPC>
F:116-191/Product: envelope protein M *status predicted <CPC>
F:130-133/Product: major envelope protein E *status predicted <NED>
F:390-733/Product: nonstructural protein NSI *status predicted <NSI>
F:101-1619/Product: hepacityrin *status predicted <NSI>
F:1214-17Region: nucleotide-binding motif A (P-loop)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        F;1620-1866/Product: nonstructural protein NS4a *status predicted <N4A>
F;1867-2017/Product: nonstructural protein NS4 *status predicted <N4B>
F;2018-3033/Product: nonstructural protein NS5 *status predicted <NS5
F;2018-303,299,305,417,423,430,448,477,534,542,558,578,627,649,1091,1217,1259,20
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1070 TQTFLGTSISGVLWTVYHGAGNKTLAGPKGPVTQMYTSAEGDLVGWPSPPGTKSLDPCTC 1129
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1010 LGREVLLGPADGYISKGWKLLAPITAYIQQTRGLLGAIVVSLIGRDKNEQAGQVQVLSSV 1069
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      53 AQTFLATCINGVCWTVYHGAGTRTIASPKGPVIQMYTNVDKDLVGWPAPQGSRSLTPCTC 112
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             10 VGRIVLNG-------AYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVSTA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      GSSDLYLVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          DB 1; Length 3033;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 65.5%; Score 668.5; DB 1;
llarity 62.6%; Pred. No. 1.8e-53;
Conservative 25; Mismatches 34;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Æ,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1.190 ARGVAKSIDFIPVESLDVATRTP 1212
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   polyprotein - marmoset hepatitis GB virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 173 TRGVAKAVDFIPVESLETTMRSP 195
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    F;1320-1323/Region: DEXH motif
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      A; Molecule type: mRNA
A; Residues: 2678-2729 <KAT>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Best Local Similarity
Matches 127; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          113
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                RESULT 12
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 음
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               A; Mcetestion: J01303
A; Molecule type: genomic RNA
A; Residues: 1-3013 < CNA
A; Cross-references: GB:D00944; NID:g221650; PIDN:BAA00792.1; PID:g221651
A; Experimental source: isolate HC-J6 from a Japanese individual
A; Experimental source: isolate HC-J6 from a Japanese individual
C; Superfamily: hepatitis C virus genome polyprotein
C; Superfamily: hepatitis C virus genome polyprotein
C; Keywords: ATP; glycoprotein; hydrolase; P-loop: polyprotein; serine proteinse; F; 2-115/Product: capsid protein M *status predicted < CNE>
F; 116-191/Product: monstructural protein E *status predicted < CNE>
F; 390-733/Product: nonstructural protein NS2 *status predicted < CNE>
F; 1010-1169/Product: nonstructural protein NS2 *status predicted < CNE>
F; 1130-1321/Region: nucleotide-binding motif B
F; 1130-1323/Region: nucleotide-binding motif B
F; 1130-1323/Redion: nonstructural protein NS4 *status predicted < CNB>
F; 1867-2017/Product: nonstructural protein NS5 *status predicted < CNB>
F; 1018-3033/Product: nonstructural protein NS5 *status predicted < CNB>
F; 1018-3033/Product: nonstructural protein NS5 *status predicted < CNB>
F; 1018-3033/Product: nonstructural protein NS5 *status predicted < CNB>
F; 1018-3033/Product: nonstructural protein NS5 *status predicted < CNB>
F; 1018-3033/Product: nonstructural protein NS5 *status predicted < CNB>
F; 1018-3033/Product: nonstructural protein NS5 *status predicted < CNB>
F; 1018-3033/Product: nonstructural protein NS5 *status predicted < CNB>
F; 1018-3033/Product: nonstructural protein NS5 *status predicted < CNB>
F; 1018-3033/Product: nonstructural protein NS5 *status predicted < CNB>
F; 1018-3033/Product: nonstructural protein NS5 *SSB, S78, 627, 649, 1091, 1217, 1259, 2038, 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      genome polyprotein - hepatitis C virus (isolate HC-J6)
N;Contains: capsid protein C; envelope protein M; hepacivirin (EC 3.4.21.98) (nonstructut protein NS4s; nonstructural protein NS5secies: hepatitis C virus
C;Species: hepatitis C virus
C;Date: 19-May-2000 *sequencc_revision 19-May-2000 *text_change 17-Nov-2030
C;Accession: JG1303 *S: Sugiyana, Y.; Kurai, K.; Tizuka, H.; Machida, A.; Miyakawa, Y. Gen Virol. 72, 2697-2704, 1991
A;Title: Nucleotide sequence of the genomic RNA of hepatitis C virus isolated from a hum A;Reference number: JG1303; MUID:92044440; PMID:1658196
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1032 AYAQQTRGVLGAIVI,SI,TGRDKNEAEGEVQFI,STATQTFI,CICINGVMTLFHGAGSKTL 1091
                                                                                                                                                                                                                  ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1092 AGPKGPVVQMYTNVDKDLVGWPSPPGKGSLTRCTCGSADLYLVTRHADVIPARRGDTRA 1151
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   53 AQTFLATCINGVCWTVYHGAGTRTIASPKGPVIQMYTNVDKDLVGWPAPQGSRSLTPCTC 112
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     113 GSSDLYI,VTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVC 172
                                                                                                                                                                                                                                                                                                                                                                                                                                                                    78 ASPKGPVIQMYINVDKDLVGWPAPQGSRSLIPCTCGSSDLYLVIRHADVIPVRRRGDSRG 137
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            -----AYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVSTA 52
                                                                                                                                                                                                                                                                                                 18 AYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVSIAAQIFLATCINGVCWTVYHGAGTRTI 77
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1152 SLLSPRPISYLKGSSGGPIMCPSGHVVGVFRAAVCTRGVAKALEFVPVENLETTMRSP 1209
                                                                                                                                                                                                                       Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      F;2015-3014/Product: nonstructural protein NS5 *status predicted <NS5> F;2210-2249/Region: interferon sensitivity determining *status predicted
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         17;
                                                                                                                                                                                                                       ;
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 DB 1: Length 3033;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Indels
                                                                                                                                                                                                                  Indels
                                                                                                                          Query Match 74.5%; Score 761; DB 1; L
Best Local Similarity 77.5%; Pred. No. 4.3c-62;
Matches 138; Conservative 19; Mismatches 21;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        62.1%; Pred. No. 1.1e-53;
Live 27; Mismatches 33
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 65.7%; Score 670.5;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1190 SRGVAKSIDFIPVETLDIVTRSP 1212
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            TRGVAKAVDFIPVESLETTMRSP 195
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            10 VGRIVLNG-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Best Local Similarity
Matches 126; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1010
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            173
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           138
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             RESULT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ρp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
```

ဟ 🗷

```
A;Residues: 1.495 <COL>
A;Cross-references: GB:AE001199; GB:AE000520; NID:g3322402; PIDN:AAC65137.1; PID:g332
A;Experimental source: strain Nichols
                                                                                                                                                                                                                                                                                                                                         Rifraser, C.M.; Norris, S.J.; Weinstock, G.M.; White, O.; Sutton, G.G.; Dodson, R rson, J.; Khalak, H.; Richardson, D.; Howell, J.K.; Chidambaram, M.; Utterback, T they, L.; Weidman, J.; Smith, H.O.; Venter, J.C. Science 281, 375-388, 1998
A:Title: Complete genome sequence of Treponema pallidum, the syphilis spirochete. A:Reference number: A71250; MUID:98332770; PMID:9665876
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           A;Status: preliminary; nucleic acid sequence not shown; translation not shown A;Wolecule type: DNA
                                                                                                                                                                                    hypothetical protein TP0136 - syphilis spirochete
C.Species: Trepomena pallidum subsp. pallidum (syphills spirochete)
C.Date: 24-Jul-1998 #sequence_revision 24-Jul-1998 #text_change 05-Nov-1999
C.Accession: B71360
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            A; Molecule type: DNA
A; Residues: 1-1334 <GLA>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          A,Gene: lin2744
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 A; Gene: TP0136
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Ω
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ð
C;Species: marmoset hepatitis GB virus A
C;Date: 23-Jul-1999 #sequence_revision 23-Jul-1969 #text_change 17-Nov-2006
C;Accession: 108839
R:Erker, J.C.; Desai, S.M.; Leary, T.P.; Chalmers, M.L.; Mostes, C.C.; Mushahwar, I.K.
J. Gen. virol. 79, 41-45, 1998
A;Title: Genomic analysis of two GB virus A variants isolated from captive monkcys.
A;Title: Genomic analysis of fwo GB virus A variants isolated from captive monkcys.
A;Title: Genomic analysis of Musb:98120818; PMID:9460920
A;Cession: T08839

                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            C.C.; Mushahwar, I.K.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               polyprotein - douroucouli hepatitis GB virus A C;Species: douroucouli hepatitis GB virus A C;Accession: T08841 A;Title: Genomic analysis of two GB virus A variants isolated from captive monkeys. A;Reference number: 216486; MUID:98120818; PMID:9460920 A;Recession: T08841 A;Reference in T08841 A;Status: translated from GB/EMBL/DDBJ A;Residues: translated from GB/EMBL/DDBJ A;Residues: 1-3005 CEBK> A;Residues: 1-3005 CEBK> A;Residues: 1-3005 CEBK> A;Cross-references: EMBL:AF023425; NID:9288599; PIDN:AAC40502.1; PID:92828600; C;Superfamily: hepatitis C virus genome polyprotein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        PCKCQPQGVWVI------RND--GALCHGTLGRTVELDLPAELCDFRGSSGSPIL 1110
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       955 RRGDEVLYGVFPSVRALPPGFVPTAPVVVMQRGIGFFSVVKTSMLGRDERFHEGSIVVIG 1014
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               346 RRGDEVLIG--VLNGVWELPPGFVPTAPVVVHHGKGFFGVVKTSMTGWDETEHVGNVVV 1003
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           51 IAAQIFLAICINGVCWIVYHGAGIRIIASPKGPVIQMYINVDKDLVGWPAPQGSRSLIPC 110
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  111 TCGSSDLYLVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAA 170
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              109 PCTCGSSDLYLVTRHADVIPVRRGDSRGSLLS------PRPISYLKGSSGGPLL 157
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             3 KKGSVVIVG----RIVLNG-----AYAQQTRGEEGCQETSQIGRDKNQVEGEVQIVS 50
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 48
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                VSTAAQTFLATCINGVCWTVYHGAGTRT1ASPKGPVIQMYTNVDKDIJGWPAPQGSRSLT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 3 KKGSVVIVGRIVLNGAYA--------QQTRGEEGCQETSQTGRDKNQVEGEVQI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Query Match 25.3%; Score 258; DB 2; Length 3005; Best Local Similarity 31.4%; Pred. No. 2.6c-15; Matches 64; Conservative 34; Mismatches 84; Indels 2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Length 2970;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CPAGHAVGIFRAAVCTRG-----VAKAVDFIPVESLETTMRSP 195
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   25.5%; Score 260; DB 2; 28.9%; Pred. No. 1.7e-15;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  38; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1132 VLHRGVKVTGVRYVKPWETLPKDS 1155
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    171 VCTRGV-----AKAVDFIPVES 187
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Local Similarity 28.9%
les 65; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    158
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   硆
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
```

```
hypothetical protein homolog lin2744 [imported] - Listeria innocua (strain Clip11262) (Species: Listeria innocua cipate 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 27-Nov-2001 (Species: Disteria innocua cipate 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 27-Nov-2001 (Species: Disterial Parche, P.; Bloec Cipate 28-Bill, AB175 (Species 28-Bill) (Spe
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           A;Cross-references: GB:AL592022; PIDN:CAC97970.1; PID:g16415280; GSPDB:GN00178
A;Experimental source: strain Clipl1262
C;Genetics:
                                                                                                                                                                                                                                                                                                                                                                                           57 LATCI -- NGVCWTVYHGAG -- - TRIIASPKGPVIQMYTNVDKDLVG -- -- - WPAPQGSR 105
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       113 VLACVPGTGVYKHCVNGAGSSSTGTTASPSTETCSQHAT----LVGGTSKPFWLVPGGTG 168
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SLTPCTC----GSSDLYLVTRHADVIP-----VRRRGDSRGSLLSPRPISYLK--- 149
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       169 NNCNCGCGGGGSSSSSSCIHIWLVPGGTGNNGNCGCGGGGSSSSSSCIHIKVEN 228
                                                                                                                                                                                                                                                                                                      3 KKGSVVIV--GRIVLNGAYAQQTRGEEGCQETSQ----TGRDKNQVEGEVQIVSTAAQTF 56
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   3 KKG---SYVIVGRIVLNGAYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVSTAAQTFLAT
                                                                                                                                                          71;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Length 1334;
               Length 495;
                                                                                                                                                   78; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         -----GSSGGPLLCPAGHAVG 165
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    229 IDEQFLDMGEGYVVTTKHLYTKNGSSSAGPAQCPGGGGG 268
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Query Match 8.4%; Score 85.5; DB 2; I
Best Local Similarity 27.0%; Pred. No. 11;
Matches 58; Conservative 19; Mismatches 83;
Query Match

8.4%; Score 85.5; DB 2;
Best Local Similarity 23.6%; Pred, No. 3.6;
Matches 52; Conservative 19; Mismatches 78;
```

738 KKGILQSLKIVDELVSVMGYPPQTIVVEMARENQTTGKGKNNSKPRYKSLEKAIKEFGSQ 797	60 CIVGWPWTYHGAGTRIIASPKGPVIQMYTNVDKDLVGWPAP 101		102 QGSRSLTPCTCGSSDLYLVTRHADVIPVRRRGDSRGSLJSPRPLS 146	847 OSFITDNSIDNLVLISSAGNREKGDDVPPLEIVRKRKVFWEKLYGGNLMSKRKFD 901	147 YL-KGSSGGPLLCPAGHAVGIFRAAVCTRGVAKAV 180	902 YLTKAERGGLTEADKARFIHRQLVETRQITKNV 934
738 KKGILQSLKIVDE	60 CI	798 ILKEHPTDNQELR	102 QGSRSLTPCTCGS	847 OSFITDNS	147 YL-KGSSGGPLLC	902 YLTKAERGGLT
<b>q</b> a	Qy	ପ୍ପ	δλ	qa	ολ	qa

Search completed: August 30, 2003, 19:20:26 Job time: 18.0488 secs

```
GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.
```

OM protein - protein search, using sw model

August 30, 2003, 18:01:52; Search time 9.65768 Seconds Run on:

(without alignments) 949.524 Million cell updates/sec

US-09-965-594-12 1021

1 MKKKGSVVIVGRIVLNGAYA......VAKAVDFIPVESLETTMRSP 195 Perfect score: Sequence:

**BLOSUM62** Scoring table:

127863 seqs, 47026705 residues Searched:

Gapop 10.0 , Capext 0.5

Total number of hits satisfying chosen parameters:

DB seq length: 0 DB seq length: 2000000000 Minimum | Maximum |

Post-processing: Minimum Match 0% Maximum Match 100% Listing first 45 summaries

SwissProt\_41:\* Database :

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

	ţi	P26664 h genome po	7958 h g	r	h genome	h genome	c	ے	h genome	treponema	arabidops	Q13685 homo sapien	-		pseudom	P08359 feline leuk				-		Q63495 rattus norv		mus mus	Q02077 feline leuk		P21414 gibbon ape			P26803 friend muri	P70289 mus musculu	P29837 1 genome po	4	Q03500 hepatitis e
SUMMARIES	ID	POLG_HCV1	POLG_HCVH	POLG_HCVTW	POLG_HCVBK	POLG_HCVJA	POLG_HCVJT	POLG_HCVJ6	POLG_HCVJB	Y136_TREPA	HHOA_ARATH	AAMP_HUMAN	DEG1_ARATH	IBP1_HUMAN	PAAD_PSEAE	ENV_FLVGL	CENE_HUMAN	PO58_CAEEL	ENV_FLVSA	ZP3_RABIT	HYPF_AZOVI	RAGE_RAT	ENV_RSVP	GRAK_MOUSE	ENV_FLVCS	INVB_DAUCA	POL_GALV	CAC3_BOVIN	ENV_FSVSM	ENV_MLVFP	PTPO_MOUSE	POLC_LANVT	GRAM_HUMAN	VST2_HEVME
	SC:	-	~	-	7	۲	-	_	-	-	-	-	-	_	-	-	-	-	-	J	-	-	-	-	~	-	-1	-1	-	-	~	_	-	-
	Length	3011	3011	3010	3010	3010	3010	3033	3033	485	321	452	437	259	209	642	2663	861	639	415	176	402	603	263	436	661	1165	253	645	919	1705	3414	257	629
ď	ch	. 1	86.1	m	7	~	2	2	2		٠	٠	7.9	7.7	7.7	7.6	7.6	7.4	7.3	7.3	•	7.2	7.2	7.2	٠	٠	٠	7.1	7.1	7.1	7.1	7.1	7.1	7.1
		894.5	879.5	847	842	838	837	670.5	668.5	85.5	82.5		80.5	79	78.5	77.5	77.5	75.5	75	74.5	74.5	74	7	73.5	73.5	ж.	•	73	7.3	73	73	73	72.5	72
	Result No.		7	3	4	2	9	7	80	6		11		13			16	17	18	19	20	21	22	23	24	25	56	27	28	29	30	31	32	33

InterPro; IPR001410; DEAD. InterPro; IPR002522; HCV\_capsid.

				Q90627 gallus gall								P18614 rattus norv		
	ENV_MLVFF	CA1H_MOUSE	GRAD_MOUSE	TRY1_CHICK	UL21_HSV11	UL21_HSV1E	OPLA_RAT	UL16_EBV	ENV_FLVLB	CN4 A_RAT	VGLB_HSVB2	ITA1_RAT	i	ALIGNMENTS
•	-1	-	г	٦	Н	~	Н	-	7	-	~			
ŗ	0/0	1527	248	248	535	535	1288	336	662	844	917	1180		
,	T · /	7.1	7.0	7.0	7.0	7.0	7.0	7.0	7.0	7.0	7.0	7.0		
ŗ	7/	72	71.5	71.5	71.5	71.5	71.5	71	71	71	71	71		
7	40	32	36	37	38	39	40	41	42	43	44	45		

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                   MEDLINE=91172826; PubMed=1848704;
MEDLINE=91172826; PubMed=1848704;
MEDLINE=91172826; PubMed=1848704;
Choo Q.-L., Richman K.H., Han J.H., Berger K., Lee C., Dong C.,
A. Callegos C., Coit D., Medina-selby A., Barr P.J., Weiner A.J.,
A. Bradley D.W., Kuo G., Houghton M.;
T. "Genetic organization and diversity of the hepatitis C virus.";
T. "Genetic organization and diversity 12455(1991).
L. Proc. Natl. Acad. Sci. U.S.A. 88:2451-2455(1991).
C. - FUNCTION: THE SMALL PROPEINS NS2A, NS2B, NS4A AND NS4B ARE
HYDROPHOBIC, SUGGESTING A POSSIBLE MEMBRANE-RELATED FUNCTION.
NS3 AND NS5 MAY PLAY A ROLE IN THE VIRAL RNA REPLICATION.
NS3 AND NS5 MAY PLAY A ROLE IN THE VIRAL RNA REPLICATION.
C. - CATALYTIC ACTIVITY: Hydrolysis of four peptide bonds in the viral position, Cys or Thr in Pl and Ser or Ala in Pl'.
C. - CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
                                                                                01-A0G-1992 (Rel. 23, Created)
01-A0G-1992 (Rel. 23, Last sequence update)
01-A0G-1992 (Rel. 23, Last sequence update)
01-S-SEP-2003 (Rel. 42, Last annotation update)
Genome polyprotein (Contains: Capsid protein C (Core protein) (P22);
Envelope glycoprotein E1 (GP32) (GP35): Envelope glycoprotein E2 (GP68) (GP70) (NG1); Protein P7: Nonstructural protein NS2 (P21)
(EC 3.4.22..); Protease/helicase NS3 (P70) (Hepacivizin)
(EC 3.4.22..98); Nonstructural protein NS34 (P4); Nonstructural protein NS48 (P27); Nonstructural protein NS58 (P66) (P70) (RMA directed RNA polymerase) (EC 2.7.7.48)].
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A LIPOPROTRIN ENVELOPE. THE ENVELOPE CONSIGTS OF TWO PROTEINS: PROTEIN AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MRNA.
                                                                                                                                                                                                                                                                                                                                                  Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY S29
                                              3011 AA.
                                              PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   EMBL; M62321; AAA45676.1; -.
                                              STANDARD:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         PIR; A39166; GNWVC3.
PDB; 1A1V; 16-FEB-99.
PDB; 1HEI; 25-NOV-98.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             MEROPS; S29.001; -. MEROPS; U39.001; -.
                                                                                                                                                                                                                                                                                                                                                                          Hepacivirus.
NCBI_TaxID=11104;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              (RNA)(N).
                                              POLG_HCV1
                        POLG_HCV1
RESULT 1
```

```
172 CTRGVAKAVDFIPVESLETTMRSP 195
                                                                                                                                                                                                                                                                                                                       PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SEQUENCE FROM N.A.
MEDLINE-92052256; PubMed-1658800;
                                                                                                                                                                                                                                                                                                                       STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Hepacivirus.
NCBI_TaxID=11108;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              (RNA)(N
                                                                                                                                                                                                                                                                                                                       POLG_HCVH
                                                                                                                                                                                                                                                                                 RESULT 2
POLG_HCVH
                                                                                     g
                                                                                                                         ò
                                                                                                                                                         a
                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ć
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             REMOVED FROM CAPSID PROTEIN C BY THE CELLULAR AMINOEPTIDASE.

CAPSID PROTEIN C (POTENTIAL).

MAJOR ENVELOPE PROTEIN B (POTENTIAL).

MAJOR ENVELOPE PROTEIN NSI/EZ (POTENTIAL).

NONSTRUCTURAL PROTEIN NSI/EZ (POTENTIAL).

PROYERASE/HELICASE NSI (POTENTIAL).

NONSTRUCTURAL PROTEIN NS4A (POTENTIAL).

NONSTRUCTURAL PROTEIN NS4A (POTENTIAL).

NONSTRUCTURAL PROTEIN NS4B (POTENTIAL).

RNA-DIRECTED RNA POLYMERASE (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Indels 11; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                               Polyprotein; Glycoprotein; Transferase; RNA-directed RNA polymerase; Core protein; Coat protein; Envelope protein; Helicase; ATP-binding; Transmembrane; Nonstructural protein; Hydrolase; Serine protease;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SIMILARITY).
SIMILARITY).
SIMILARITY).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            (POTENTIAL)
(POTENTIAL)
(POTENTIAL).
(POTENTIAL).
(POTENTIAL).
(POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       87.6%; Score 894.5; DB 1; Length 3011;
85.8%; Pred. No. 2.2e-76;
tive 8; Mismatches 10; Indeis 11;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     65F8C9447FCE5AF9 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CHARGE RELAY SYSTEM (BY
CHARGE RELAY SYSTEM (BY
CHARGE RELAY SYSTEM (BY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (GLCNAC. . .
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              (GLCNAC. .
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      N-LINKED (GLCNAC. IN-LINKED (GLCNAC. N-LINKED GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ATP (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   POTENTIAL.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      N-LINKED (N-LINKED (N-LINK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          DECH BOX
                                                                                                                       rpro: IPR002868; NCV_NSSa.
rpro: IPR002166; HCV_RGARP.
rpro: IPR001659; Helicase_C.
rpro: IPR070995; RNA_pol_bS_PS.
rpro: IPR007094; RNA_pol_PSvir.
PP01543; HCV_capsid; 1.
                                                                                                                                                                                                                  327197 MW;
                                                                    HCV_NS3.
HCV_NS4a.
HCV_NS4b.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               369
1083
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                2240
2364
2789
                                                                                   IPRO00745; 1
IPR001490; 1
IPR002868; 1
                                IPR002531;
IPR002518;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Š
                                                                        IPR004109;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Best Local Similarity
Matches 175; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    3D-structure.
INIT_MET
                                                                                                        InterPro; I
                                                                      InterPro;
InterPro;
               InterPro;
InterPro;
InterPro;
                                                                                                                                          InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ACT_SITE
ACT_SITE
NP_BIND
SITE
                                                                                                                                                             InterPro;
                                                                                                                                                                                             InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 TRANSMEM
ACT_SITE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CHAIN
CHAIN
CHAIN
CHAIN
CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CHAIN
```

51

----AYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVST

KKGSVVIVG - - - RIVLNG - - -

```
1005 RRGREILLGPADGMVSKGWRLLAPITAYAQQTRGLLGCIITSLTGRDKNQVEGEVQIVST 1064
                                                                          01-AUG-1992 (Rel. 23, Created)
01-AUG-1992 (Rel. 23, Last sequence update)
01-AUG-1992 (Rel. 24, Last annotation update)
01-S-EP-2003 (Rel. 42, Last annotation update)
Genome polyprotein [Contains: Capsid protein C (Core protein (P22);
Envelope glycoprotein El (GP32) (GP35); Envelope glycoprotein E2
(GP68) (GP70) (NS1); Protein P7; Nonstructural protein NS2 (P21)
(EC 3.4.99-.); Protease/helicase NS3 (P70) (Hepacivirin)
(EC 3.4.21.98); Nonstructural protein NS4A (P4); Nonstructural protein
NS4B (P72); Nonstructural protein NS5A (P56); Nonstructural protein
NS5B (P60) (RNA-directed RNA polymerase) (EC 2.7.7.48)].
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            X-RAY CRYSTALLOGRAPHY (2.1 ANGSTROMS) OF 1207-1657.
MEDLINE-97331322; PubMed-9187654;
Yao N., Hesson T., Cable M., Hong Z., Kwong A.D., Le H.V., Weber P.C.;
Structure of the hepatitis C virus RNA helicase domain.";
Nat. Struct. Biol. 4:463-467(1997).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS) OF 1192-1657.
MEDLINE-98154321; PubMed-9493270;
Kim J.L., Morgenstern K.A., Griffith J.P., Dwyer M.D., Thomson J.A.,
Murcko M.A., Lin C., Caron P.R.;
"Hepatitis C virus NS3 RNA helicase domain with a bound
oligonucleotide: the crystal structure provides insights into the mode
                                                52 AAQTFLATCINGVÇWTVYHGAGTRTIASPKGPVIQMYTNVDKDLVGWPAPQGSRSLTPCT
                                                                                                                                             112 CGSSDLYLVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      of unwinding.";
structure 6:89-100(1998).
-!- FUNCTION: PROTENSE NS IS RESPONSIBLE FOR THE CLEAVAGE OF NS2-NS3.
-!- FUNCTION: PROTENSE NS IS RESPONSIBLE FOR THE CLEAVAGE OF
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          -:- FUNCTION: NSSA SEEMS TO HAVE A TRANSCRIPTIONAL ACTIVATORY ROLE.
-:- FUNCTION: NSSB IS A RNA-DEPENDENT RNA POLYMERASE THAT PLAYS AN
FSSEMTIAL ROLE IN THE VIROS REPLICATION.
-:- CATALYTIC ACTIVITY: Hydrolysis of four peptide bonds in the viral
precursor polyprotein, commonly with Asp or Glu in the P6
position, Cys or Thr in P1 and Ser or Ala in P1'.
-:- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          NS3-NS4A, NS4A-NS4B, NS4B-NS5A AND NS5A-NS5B.
-!- FUNCTION: NS4A FORMS A COMPLEX WITH NS3 AND IS ESSENTIAL FOR THE ACTIVATION OF NS3.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS: E1 AND E2. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MRNA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     hepatitis
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Inchauspe G., Zebedee S., Lee D.H.H., Sugitani M., Nasoff M.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Prince A.M.;

"Genomic structure of the human prototype strain H of
"Yeronic structure of the human prototype strain H of
virus: comparison with American and Japanese isolates.

Proc. Natl. Acad. Sci. U.S.A. 88:10292-10296(1991).
                                                                                                                                                                                                                                                                                                                                                                                                             3011 AA
```

```
(POTENTIAL)
    (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Length 3011;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              327142 MW; 772CBB29CCD94753 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Score 879.5; DB 1;
Pred. No. 5.9e-75;
  (GLCNAC.
  N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    86.1%;
83.8%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1417
1420
1420
1436
1448
1478
1488
11480
11501
                                                                                                                              645

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1226

1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ¥.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Query Match
Best Local Similarity
   1511
1514
1532
1533
1555
1555
1566
1566
1667
1627
1635
  CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
                                                         CARBOHYD
CARBOHYD
CARBOHYD
                                                                                           CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
TURN
TURN
TURN
STRAND
TURN
STRAND
HELIX
TURN
STRAND
STRAND
STRAND
HELIX
TURN
STRAND
STRAND
HELIX
TURN
TURN
STRAND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                HELIX
   This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the Buropean Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                  SIMILARITY).
SIMILARITY).
SIMILARITY).
PTM: THE STRUCTURAL PROTEINS C, E1 AND E2 ARE PRODUCED BY PROTEOLYTIC PROCESSING BY THE HOST SIGNAL PEPPIDASES. SIMILARITY: THE NS2 PROTEASE BELONGS TO PERPIDASE FAMILY US9. SIMILARITY: THE NS3 PROTEASE BELONGS TO PEPTIDASE FAMILY S29.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     LINKED (GLCNAC. .) (POTENTIAL).
LINKED (GLCNAC. .) (POTENTIAL).
LINKED (GLCNAC. .) (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           REMOVED FROM CAPSID PROTEIN C
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CELLULAR AMINOPEPTIDASE.
CAPSID PROTEIN C.
ENVELOPE GLYCOPROTEIN E1.
ENVELOPE GLYCOPROTEIN E1.
PROTEIN P7.
NONSTRUCTURAL PROTEIN NS2.
PROTEASE/HELICASE NS3.
NONSTRUCTURAL PROTEIN NS4A.
NONSTRUCTURAL PROTEIN NS4A.
NONSTRUCTURAL PROTEIN NS5A.
NONSTRUCTURAL PROTEIN NS5A.
NONSTRUCTURAL PROTEIN NS5A.
NONSTRUCTURAL PROTEIN NS5A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              (BY
(BY
(3Y
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  POTENTIAL.
CHARGE RELAY SYSTEM (1
CHARGE RELAY SYSTEM (1
CHARGE RELAY SYSTEM (1)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ATP (POTENTIAL).
DECH BOX.
                                                                                                                                                  EMBL; M67463; AAA45534.1; ~.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  A36814; GNWVCH.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           3D-structure.
INIT_MET
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CARBOHYD
CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               TRANSMEM
ACT_SITE
ACT_SITE
ACT_SITE
NP_BIND
SITE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CHAIN
CHAIN
CHAIN
CHAIN
CHAIN
CHAIN
CHAIN
CHAIN
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bloinformatics and the EMBL outstation between Bloinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/
                                                                   52 AAQTFLATCINGVCWTVYHGAGTRTIASPKGPVIQMYINVDKDLVGWPAPQGSRSLTPCT 111
                                                                                                                                                                                                         112 CGSSDLYLVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPILCPAGHAVG1FRAAV 171
                                           3 KKGSVVIVG---RIVLNG----AYAQOTRGEEGCQFTSQTGRDKNQVEGEVQIVST 51
                                                                                                                                                                                                                                                                                                                                                                                                                                                                    01-APR-1993 (Rel. 25, Created)
01-APR-1993 (Rel. 25, Last sequence update)
15-SEP-2003 (Rel. 42, Last annocation update)
6enome polyprotein (Contains: Capsid protein C (Core protein) (P22);
6enome polyprotein El (GP32) (GP35); Envelope glycoprotein E2 (GP68) (GP70) (NSI); Protein P7; Nonstructural protein NS2 (P21) (EC 3.4.22 -); Protease/helicase NS3 (P70) (Hepacivirin) (EC 3.4.21.98); Nonstructural protein NS4A (P4); Nonstructural protein NS5A (P56); Nonstructural protein NS5A (P66) (P70) (RNA-directed RNA polymerase) (EC 2.7.7.48)].
Hepathis C virus (Isolate Taiwan) (HCV).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          (RNA)(N).
SUBDNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MRNA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY $29.
  13; Indels
  9; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                            1185 CTRGVAKAVDFIPVENLETTMRSF 1208
                                                                                                                                                                                                                                                                                         172 CTRGVAKAVDFIPVESLETTMRSP 195
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 EMBL; M84754; -; NOT_ANNOTATED_CDS
PIR; A40244; GNWVIW.
Matches 171; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                   STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        PDB; 1N64; 25-FEB-03.
PDB; 1NS3; 08-APR-98.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 MEROPS; $29.001; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Hepacivirus.
NCBI_TaxID-31645;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         039.00
                                                                                                                                                                                                                                                                                                                                                                                                                                   POLG_HCVTW
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         MEROPS;
                                                                                                                                                                                                                                                                                                                                                                                                                                   THE STATE OF STATE AND STA
                                                                                                                           ð
                                                                                                                                                             qq
                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                         a
```

InterPro; IPR001410; DEAD.

```
ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CELLUIAR AMINOPEPTIDASE.

CELLUIAR AMINOPEPTIDASE.

MATRIX PROTEIN (POTENTIAL).

MAJOR ENVELOPE PROTEIN E (POTENTIAL).

MAJOR ENVELOPE PROTEIN NSI/E2 (POTENTIAL).

NONSTRUCTURAL PROTEIN NS2 (POTENTIAL).

PROTEASE/HELICASE NS3 (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Probom; PD186062; HCV_NSI; 1.
SMART; SM00487; DEXDC; 1.
Polyprotein; Glycoprotein; Transferase; RNA-directed RNA polymerase;
Core protein; Coat protein; Envelope protein; Helicase; ATP-binding;
Transmembrane; Nonstructural protein; Hydrolase; Serine protease;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          NONSTRUCTURAL PROTEIN NS4A (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4B (POTENTIAL).
RNA-DIRECTED RNA POLYMERASE (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      REMOVED FROM CAPSID PROTEIN C BY THE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CHARGE RELAY SYSTEM (BY SIMILARITY)
CHARGE RELAY SYSTEM (BY SIMILARITY)
CHARGE RELAY SYSTEM (BY SIMILARITY)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Length 3010;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         MW; AAD267D55CDFE215 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (GLCNAC. . ) (GLCNAC. . ) (GLCNAC. . ) (GLCNAC. . ) (
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (GLCNAC. . .)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              (GLCNAC. . .)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (GLCNAC. ..)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (GLCNAC. . .)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (GLCNAC. . .
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              83.0%; Score 847; DB 1;
88.2%; Pred. No. 7.1e-72;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          N-LINKED (GLCNAC.
N-LINKED (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     N-LINKED (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  (GICNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ATP (POTENTIAL)
DECH BOX.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   N-LINKED (
N-LINKED (
N-LINKED (
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            POTENTIAL.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 N-LINKED (N-LINKED (N-LINK
                                                                                                                                                                                                                                        Pro; IPR007095; RNA_pol_DS_PS.
Pro; IPR007094; RNA_pol_PSvir.
PF01543; HCV_capsid; 1.
PF01542; HCV_core; 1.
                                                                                                                                                                                                                                                                                                                                             Pfam; PF01539; HCV_env; 1.
Pfam; PF01569; HCV_env; 1.
Pfam; PF01560; HCV_NS1; 1.
Pfam; PF01006; HCV_NS3; 1.
Pfam; PF01006; HCV_NS4; 1.
Pfam; PF01001; HCV_NS4; 1.
Pfam; PF01506; HCV_NS5; 1.
Pfam; PF00271; HCV_NS5; 1.
Pfam; PF0099; Viral_RGRP; 1.
ProDom; PP186062; HCV_NS1; 1.
HCV_COTE
HCV_ENS.
HCV_NS2.
HCV_NS3.
HCV_NS3.
HCV_NS4a.
HCV_NS5a.
HCV_NS5a.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         327047
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        115
191
383
729
1006
11615
1862
2013
369
369
                                                                                                                                                                                                                       IPR002166;
                                                                                                                                                                                                                                              interPro; IPR007095;
InterPro: IPR007094;
                                                                                                                                      IPR000745;
                                                                                                                                                                  InterPro; IPR001490;
                                                                                                                                                                                           IPR002868;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Query Match
Best Local Similarity
Matches 157; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               3D-structure.
INIT_MET
                                                                                                             InterPro;
InterPro;
                                                                                                                                                                                              interPro;
                                                                                                                                                                                                                  InterPro;
                                                                                                                                                                                                                                                                       InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ACT_SITE
NP_BIND
SITE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            LRANSMEM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ACI_SITE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CHAIN
CHAIN
CHAIN
                                                                                                                                                                                                                                                                                                Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CHAIN
                                                                                                                                                                                                                                                                                                                            Pfam;
```

```
PROTEIN C AND MRNA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         FRANSMEM
                                                                                                                                                                                                                                                                                                                                                                                                                                               MEROPS;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CHAIN
CHAIN
CHAIN
CHAIN
CHAIN
CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CHAIN
1091 AGPKGPITQMYTNVDQDLVGWHAPQGARSLTPCTCGSSDLYLVTRHADVIPVRRRGDSRG 1150
                                    1031 AYAQQTRG1.FGCITTSLTGRDKNQVEGEVQVVSTATQSF1.ATCINGVCWTVYHGAGSKTL 1090
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             X **RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS) OF 1027-1210 AND 1678-1691.

X MEDLINE-98227846; PubMed-9568891;
Yan Y., Li Y., Munshi S., Sardana W.,
Steinkuehler C., Tomei L., de Francesco R., Kuo I.C., Chen Z.;
T 'Complex of NS3 protease and NS4A peptide of BK strain hepatitis C
T virus: a 2.2-A resolution structure in a hexagonal crystal form.";
Protein Sci. 7:837-847(1998).

C Protein Sci. 7:837-847(1998).

- :- FUNCTION: THE SMALL PROTEINS NS2B, NS4B, AND NS4B ARE
HYDROPHOBIC, SUGGESTING A POSSIBLE MEMBRANE-RELATED FUNCTION.
NS3 AND NS5 MAY PLAY A ROLE IN THE VIRAL RNA REPLICATION.

C CATALYTIC ACTIVITY: Hydrolysis of four peptide bonds in the viral presention, Cys or Thr in Pl and Ser or Ala in Pl'.

C CATALYTIC ACTIVITY: N nucleoside triphosphate ~ N diphosphate +
                                                                            ASPKGPVIQMYTNVDKDLVGWPAPQGSRSLTPCTCGSSDLYLVTRHADVIPVRRKGDSRG 137
AYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVSTAAQTFLATCINGVCWTVYHGAGTRTI 77
                                                                                                                                                                                 SILSPRPISYLKGSSGGPLICPAGHAVGIFRAAVCTRGVAKAVDFIPVESLETTMRSP 195
                                                                                                                                                                                                                                                                                                                                                       11-5EP-2003 (Rel. 23, Last annotation update)
15-5EP-2003 (Rel. 42, Last annotation update)
16-5EP-2003 (Rel. 42, Last annotation update)
16-5EP-2003 (Rel. 42, Last annotation update)
16-5ED-2003 (Rel. 42, Last annotation update)
16-5ED-2003 (Rel. 42, Last annotation update)
16-5ED-2003 (Rel. 42, Last annotation Elegan (GP48) (GP38); Envelope glycoprotein Elecation Elecation Elecation Elecation Elecation Elecation (Rel. 42, Last); Protesin Pl.; Nonstructural protein NSA (Ph): Nonstructural Protein NSA (Ph
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A LIPOPROTEIN ENVELOPE CONSISTS OF TWO PROTEINS: PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Takamizawa A., Mori C., Fukc I., Manabo S., Murakami S., Fujita J., Onishi E., Andoh T., Yoshida I., Okayama H.;
"Structure and organization of the hepatitis C virus genome isolated
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ď
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   X-RAY CRYSTALLOGRAPHY (2.4 ANGSTROMS) OF 1027-:215.

BUDILNE-PROT16808; Pubmed-8861916;

Love R.A., Parge H.E., Wickersham J.A., Hostomsky Z., Habuka N., Moomaw E.W., Adachi T., Mostomska Z.;

The crystal structure of hepatitis C virus NS3 proteinase reveals trypsin-like joid and a structural zinc binding site.";

Cell 87:331-342(1996).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                MEDIINE-96235224; PubMed-8647104;
Borcowski P., Heiland M., Oebhmann K., Becker B., Korneteky L.;
"Non-structural protein 3 of hepatitis C virus inhibits
phosphorylation mediated by cAMP-dependent protein kinase.";
Eur. J. Riochem. 237:611-618(1996).
                                                                                                                                                                                                                                                                                                      PRT; 3010 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SEQUENCE FROM N.A.
MEDLINE-91140698; Pubmed-1847440;
                                                                                                                                                                                                                                                                                                                                               (Rel. 23, Created)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Virol. 65:1105-1113(1991).
                                                                                                                                                                                                                                                                                                      STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SEQUENCE OF 1487-1500
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   from human carriers.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        NCBI_TaxID=11105;
                                                                                                                                                                                                                                                                                                                                          01-AUG-1992
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Hepacivirus
                                                                                                                                                                                                                                                                                                      POLG_HCVBK
P26663;
                                                                                                                                                           138
                                                                            7.8
18
                                                                                                                                                                                                                                                         q
                                                                            ò
                                                                                                                 g
                                                                                                                                                         ò
                                                                                                                                                                                                  a
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation. The European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CELLULAR AMINOPEPTIDASE.
CAPSID PROTEIN C (POTENTIAL).
MATRIX PROTEIN (POTENTIAL).
MAJOR ENVELOPE PROTEIN E (POTENTIAL).
NONSTRUCTURAL PROTEIN NS1,EZ (POTENTIAL).
PRONTRUCTURAL PROTEIN NS2 (POTENTIAL).
PROTEASE/HELICASE NS3 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS48 (POTENTIAL).
RNA-DIRECTED RNA POLYMERASE (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Polyprotein; Glycoprotein; Transferase; RNA-directed RNA polymerase; Core protein; Coat protein; Envelope protein; Helicase; ArP-binding; Transmembrane; Nonstructural protein; Hydrolase; Serine protease;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   REMOVED FROM CAPSID PROTEIN C BY THE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            N-LINKED (GLCNAC. ..) (POTENTIAL).
N-LINKED (GLCNAC. ..) (POTENTIAL).
N-LINKED (GLCNAC. ..) (POTENTIAL).
SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY S29
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CHARGE RELAY SYSTEM.
CHARGE RELAY SYSTEM.
CHARGE RELAY SYSTEM.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ATP (POTENTIAL).
DECH BOX.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            POTENTIAL.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               InterPro; IPR001410; DEAD.
InterPro; IPR001219; HCv_capsid.
InterPro; IPR001219; HCv_care.
InterPro; IPR001219; HCv_care.
InterPro; IPR001219; HCv_care.
InterPro; IPR001219; HCv_MS1.
InterPro; IPR0014109; HCv_MS3.
InterPro; IPR001409; HCv_MS4.
InterPro; IPR001409; HCv_MS4.
InterPro; IPR001409; HCv_MS4.
InterPro; IPR001868; HCv_MS4.
InterPro; IPR001868; HCv_MS4.
InterPro; IPR001868; HCv_MS7.
InterPro; IPR001995; RNA_pol_DS.
InterPro; IPR001995; RNA_pol_DS.
InterPro; IPR001995; HOV_COPP.
IPR0M: PF001598; HCV_MS1; I.
IPR0M: PF001598; HCV_MS1; I.
IPR0M: PF001598; HCV_MS2; I.
IPR0M: PF001001; HCV_MS49; I.
IPR0M: PF001001; HCV_MS49; I.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Pfam; PF01506; HCV_NS5a; 1.
Pfam; PF00998; Viral_RdRP; 1.
ProDom; PD186062; HCV_NS1; 1.
                                                                                                                                                                                                                                                 EMBL; M58335; AAA72945.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1615
1862
2013
3010
369
1083
11065
11165
1237
209
209
234
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SMART; SM00487; DEXDC;
                                                                                                                                                                                                                                                                                             1A1Q; 25-MAR-98.
1JXP; 14-JAN-98.
1C2P; 15-NOV-00.
1C2J; 08-NOV-99.
1GX5; 09-APR-02.
1GX6; 10-APR-02.
1GX6; 10-APR-02.
                                                                                                                                                                                                                                                                           GNWVTC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          347
1083
1107
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1165
1230
1316
196
209
234
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               3D-structure.
ENII_MET
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        $29.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               U39.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ACT_SITE
ACT_SITE
ACT_SITE
NP_BIND
SITE
CARBOHYD
CARBOHYD
```

ø

```
MEROPS; U39.00
                                                                                                                                                                                                                                                                                                                                                                                     InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                   InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                              interPro;
                                                                                                                                                                                                                                                                                                                                                                                                                     InterPro;
                                                                                                                                                                                                                                                                                                                                                                        InterPro;
                                                                                                                                                                                                                                                                                                                                                                                InterPro;
                                                                                                                                                                                                                                                                                                                                                                                            InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                        InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                            InterPro;
                                                                                                                                                                                                                                                                                                                                                                                          0
                                                                                                                                                                                                                                                                                                                                                                                   78 ASPKGPVIQMYTNVDKDLVGWPAPQGSRSLTPCTCGSSDLYLVTRHADVIPVRRRGDSRG 137
                                                                                                                                                                                                                                                                                                                                                          18 AYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVSTAAQTFLATCINGVCWIVYHGAGTRTI 77
                                                                                                                                                                                                                                                                                                                                                                                                                   SLLSPRPISYLKGSSGGPLL.CPAGHAVGIFKAAVCTRCVAKAVDFIPVESLETTMRSP 195
                                                                                                                                                                                                                                                                                                                                               Gaps
                         (POTENTIAL)
(POTENTIAL)
(POTENTIAL)
(POTENTIAL)
(POTENTIAL)
(POTENTIAL)
(POTENTIAL)
      (POTENTIAL).
(POTENTIAL).
(POTENTIAL).
                                                                      (POTENTIAL)
                                                                                               (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                               .;
                                                                                   (POTENTIAL)
                                                                                          (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                 82.5%; Score 842; DB 1; Length 3010;
86.5%; Pred. No. 2.1e-71;
ive 14; Mismatches 10; Indels
                                                                                                                                                                                                                                                                                                                     327189 MW; F8422D5ECCFDFD9C CRC64;
                                                                                        (GLCNAC.
(GLCNAC.
(GLCNAC.
                         (GLCNAC.
(GLCNAC.
(GLCNAC.
(GLCNAC.
(GLCNAC.
                                                                     (GLCNAC.
(GLCNAC.
             (GLCNAC
        (GLCNAC
                                                               (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                          3010 A.A
N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                         POLG HCVJA STANDARD;
P26662;
01-AUG-1992 (Rel. 23, Created)
                                                                                                                                                                                                                                                                                                                                        Best Local Similarity 86.59
Matches 154; Conservative
                                                                                                                                                                                                                                                                                     11188
11197
1202
1204
                                                                                        2240
2529
2529
2788
1035
1004
1004
1006
1008
1009
1009
11003
                                                                                                                                                                                                                                                                                                                      Ą.
                                                                                                            10031
10059
10059
10059
10059
10062
10062
10062
10062
10063
10063
10063
10063
10063
10063
10063
10063
10063
10063
10063
10063
      CARBOHYD
CARBOHYD
CARBOHYD
                                                        CARBOHYD
CARBOHYD
CARBOHYD
                                                                          CARBOHYD
CARBOHYD
CARBOHYD
                                                                                                     CARBOHYD
STRAND
                                                                                                                                                                                                                                                                                                                                                                                                             138
                                                                                                                                                                                                                                                                                                                                                                                                                          1151
                                                                                                                                                                                                                                                                                                                      SEQUENCE
                                                                                                                                                                                                                                                                                                                                 Query Match
                           CARBOHYD
                                CARBOHYD
                                      CARBOHYD
                                             CARBOHYD
                                                   CARBOHYD
                                                                                               CARBOHYD
                                                                                                                              STRAND
TURN
STRAND
HELIX
HELIX
TURN
TURN
TURN
STRAND
STRAND
STRAND
STRAND
STRAND
STRAND
                                                                                                                                                                                                                                  STRAND
STRAND
HELIX
TORN
                                                                                                                                                                                                                                                                  STRAND
TURN
STRAND
                                                                                                                 HELIX
STRAND
                                                                                                                                                                                                                                                                                                                                                                                                                                                   POLG_HCVJA
ID POLG_HC
AC P26662
DT 01-AGG
                                                                                                                                                                                                                                                                                           STRAND
                                                                                                                                                                                                                                                                                                               STRAND
                                                                                                                                                                                                                                                                                     TURN
                                                                                                                                                                                                                                                                                                                                                                                                                                             RESULT
ò
                                                                                                                                                                                                                                                                                                                                                                     q
                                                                                                                                                                                                                                                                                                                                                                                   õ
                                                                                                                                                                                                                                                                                                                                                                                               a
                                                                                                                                                                                                                                                                                                                                                                                                            õ
                                                                                                                                                                                                                                                                                                                                                                                                                        g
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its wee by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce.or send an email to license@isb-sib.ch).
01-AUG-1992 (Rel. 23, Last sequence update)
28-FEB-2003 (Rel. 41, Last annotation update)
6enome polyprotein [Contains: Capsid protein C (Core protein) (P22);
6ENSALORO (NS1); Protein El (GP22) (GP35); Envelope glycoprotein El (GP69) (GP70) (NS1); Protein P7; Nonstructural protein NS2 (P21)
6EC 3.4.22.-); Proteasse/helicase NS3 (P70) (Hepacivirin)
6EC 3.4.22.-); Proteasse/helicase NS3 (P70) (Hepacivirin)
8EQ 9E27); Nonstructural protein NS5A (P4); Nonstructural protein NS5B (P56); P70) (RNA-directed RNA polymerase) (EC 2.7.7.48)).
8ES P66) (P70) (RNA-directed RNA polymerase) (EC 2.7.7.48)).
8ES RNA POSITIVE-Strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                                                                       MEDINE-91192166; PubMed-1849488;

Kato N., Hijikata M., Nakagawa M., Ootsuyama Y., Muraiso K.,
Ohkashi S., Shimchoho K.;
Ohkashi S., Shimchoho K.;
Wolecular structure of the Japanese hepatitis C viral genome.";
FEBS Lett. 280:325-328(1991).
-!-PONTION: THE SYALL PROTEINS NS2A, NS2B, NS4A AND NS4B ARE
HYDROPHOBIC, SUGGESTING A POSSIBLE MEMBRANE-RELATED FUNCTION.
NS3 AND NS5 MAY PLAY A ROLE IN THE VIRAL RNA REPLICATION.
-!-CATALYTIC ACTIVITY: Hydrolysis of four peptide bonds in the viral
precursor polyprotein, commonly with Asp or Glu in the P6
position, Cys or Thr in Pl and Ser or Ala in Pl'.
-!-CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           {RNA}(N).
-1- SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLECCAPSID COVERED BY LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
PROTEIN M AND GLYCOPROTEIN E. THE NUCLECCAPSID IS A COMPLEX OF PROTEIN C AND MRNA.
                                                                                                                                                                                                                                                                    SEQUENCE FROM N.A.
MEDLINE-91088550; Pubmed-2175903;
Kato N., Hilkata M., Ootsuyama Y., Nakagawa M., Ohkoshi S.,
Sugimura T., Shimotohno K.;
"Molecular cloning of the human hepatitis C virus genome from
"Japanese patients with non-8 hepatitiss";
Proc. Natl. Acad. Sci. U.S.A. 87:9524-9528(1990).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       -!- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY S29
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      InterPro; IPR001650; Helicase_C.
InterPro; IPR007095; RNA_pol_DS_PS.
InterPro; IPR007094; RNA_pol_PSv1r.
Pfam; PF01543; HCV_core: 1.
Pfam; PF01539; HCV_core: 1.
Pfam; PF01539; HCV_env: 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                HCV_capsid.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       HCV_NS1.
HCV_NS2.
HCV_NS3.
HCV_NS4a.
HCV_NS4b.
HCV_NS5a.
HCV_NS5a.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     EMBL; D90208; BAA14233.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                             DISCUSSION OF SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       IPR002531; H
IPR002518; H
IPR004109; H
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          IPR002166;
IPR001650;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   PIR; A39253; GNWVCJ.
HSSP; P26663; lJXP.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              InterPro; IPR001410;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   InterPro; IPR002522;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         IPR002519;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   IPR001490;
IPR002868;
                                                                                                                                                                                                                                  NCBI_TaxID=11116;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         MEROPS; S29.001;
```

```
MEROPS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Pfam; P
Pfam; P
Pfam; P
fam: PF01538; HCV_NS2: 1.

Pfam: PF02907; HCV_NS3; 1.

Pfam: PF01001; HCV_NS4; 1.

Pfam: PF01001; HCV_NS4; 1.

Pfam: PF01001; HCV_NS4; 1.

Pfam: PF01003; HCV_NS4; 1.

Pfam: PF01003; HCV_NS4; 1.

Pfam: PF01003; HCV_NS4; 1.

Pfam: PF01003; HCV_NS4; 1.

Pfam: PF01099; Viral_RdRP; 1.

NART: SM00487; DEXDC: 1.

NONSTRUCTURAL PROTEIN NS! (POTENTIAL).

NONSTRUCTURA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              78 ASPKGPVIQMYINVDKDLVGWPAPQGSRSLTPCTCGSSDLYLVTRHADVIPVRRGDSRG 137
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               77
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVCTRGVAKAVDFIPVESLETIMRSP 195
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               18 AYAQQIRGEEGCQETSQIGRDKNQVEGEVQIVSTAAQIFILATCINGVCWTVYHGAGTRII
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    0; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Score 838; DB 1; Length 3010;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         85.4%; Pred. No. 5.1e-71;
tive 16; Mismatches 10; Indeis
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     AA993794F46DB185 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 N-LINED (GLCNAC,
N-LINED (GLCNAC,
N-LINED GLCNAC,
N-LINED GLCN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Created)
Last sequence update)
Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     AA; 327017 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                82.18;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Best Local Similarity 85.4 Matches 152; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       645
2041
2077
2240
2788
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (Rel. 25, (Rel. 25, I (Rel. 25, I
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               RESULT 6
POLG_HCVJT
ID POLG HCVJT
AC Q00269;
DT 01-APR-1993 (
DT 01-APR-1993 (
DT 15-SEP-2003 (
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CARBOHYD
                                                                                                                        a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMED custration the European Bloinformatics institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@lsb-sib.ch).
                                                                                                                                                                                                                                                                                                           MEDLINE-92295714; PubMed-1318627;
MEDLINE-92295714; PubMed-1318627;
A Tanaka T., Kato N., Nakagawa M., Ootsuyama Y., Cho M.J.,
Nakazawa T., Hijikata W., Ishimura Y., Shimotohno K.;
Molecular cloning of hepatitis C virus genome from a single Japanese carier: sequence variation within the same individual and among infected individuals.";
Virus Res. 23:39-53(1992).
L. PUNCTRON: THE SMALL PROTEINS NSZA, NSZB, NS4A AND NS4B ARE
I- PUNCTRON: THE SMALL PROTEINS NSZA, NSZB, NS4B AND NSS AN
Genome polyprotein [Contains: Capsid protein C (Core protein) (P22); Envelope glycoprotein E1 (GP32) (GP35); Envelope glycoprotein E2 (GP68) (GP70) (NS1); Protein P7: Nonstructural protein NS2 (P21) (E2 3.4.22.-); Protease/helicase NS3 (P70) (Heperivirin) (EC 3.4.21.98); Nonstructural protein NS48 (P4); Nonstructural protein NS54 (P56); Nonstructural protein NS54 (P56), Nonstructural protein NS54 (P65); Nonstructural protein NS54 (P60) (P70) (NA-directed NA polymerase) (EC 2.7.7.48)]. Hepatitis C virus (isolate HC-JT) (HCV).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             -1. SUBGNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS: PROTEIN M AND GLYCOPROTEIN E, THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MRNA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              -!- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY S29.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      InterPro: IPR001410; DEAD.
InterPro: IPR002522; HCV_capsid.
InterPro: IPR002521; HCV_core.
InterPro: IPR002519; HCV_env.
InterPro: IPR002519; HCV_NSI.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     EMBL; D11168; BAA01943.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           helicase_C; 1
Viral_RdRP; 1
62; HCV_NS1; 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              PIR; A45573; A45573;
PDB; 1A1Q; 25-MAR-98.
PDB; 1JXP; 14-JAN-98.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         PF00998; Vir
i; PD186062;
                                                                                                                                                                                                                                                                  NCBI_TaxID=31642;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              MEROPS; S29.001;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         039.00
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            (RNA)(N).
                                                                                                                                                                                                                                            Hepaciv1rus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Pfam; PF(
Probom; 1
```

œ

```
Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Pfam;
1031 AYAQQTRGILGCIVTSLTGRDKNQVEGEVOVVSTATQSFLATCVNGVCWTVFHGAGSKTL 1090
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            0;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ASPKGPVIQMYTNVDKDLVGWPAPQGSRSLTPCTCGSSDLYI,VTRHADVIPVRRRGDSRG 137
                                                                                                                                                                                                                                                                                                                                                                                                                                                              11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CELLULAR AMINOPEPIDASE.
CAPSID PROTEIN C (POTENTIAL).
MATRIX PROTEIN (POTENTIAL).
MACON REVELOPE PROTEIN E (POTENTIAL).
NONSTRUCTURAL PROTEIN NS! / ROTENTIAL).
NON-STRUCTURAL PROTEIN NS2 (POTENTIAL).
PROTEASE/HELICASE NS3 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4A (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4A (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4B (POTENTIAL).
RNA-DIRECTED RNA POLYMERASE (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SLLSPRPISYLKGSSGGPLLCPAGHAVG1FRAAVCTRGVAKAVDF1PVESLETTMRSP 195
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           01-A0G-1992 (Rel. 23, Created)
01-M0G-1992 (Rel. 23, Last sequence update)
28-FEB-2003 (Rel. 41, Last annotation update)
28-FEB-2013 (Rel. 41, Last annotation update)
Genome polyprotein (Contains: Capsid protein C (Core protein) (P22);
Envelope glycoprotein El (GP32) (GP35): Envelope glycoprotein E2
(GP68) (GP70) (NS1): Protein P7: Nonstructural protein NS2 (P21)
(EC 3.4.22.-): Protease/Nelicase NS3 (P70) (Bepaciation)
(EC 3.4.21.98); Nonstructural protein NS4A (P4); Nonstructural protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                             0; Gaps
SNART; SM00487; DEXDC; 1.
Subsprotelain, Glycoprotelain, Transferase; RNA-directed RNA polymerase;
Core protein; Coat protein; Envelope protein; Helicase; ATP-binding:
Transmembrane; Nonstructural protein; Hydrolase; Serine protese;
                                                                                                                                                              SIMILARITY).
SIMILARITY).
SIMILARITY).
                                               REMOVED FROM CAPSID PROTEIN C BY THE
                                                                                                                                                                                                                                                                                                                                    (POTENTIAL)
                                                                                                                                                                                                                                                 (POTENTIAL)
                                                                                                                                                                                                                                                                                                                           (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                 (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                   (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                            POTENTIAL)
                                                                                                                                                                                                                                                                                     (POTENTIAL)
                                                                                                                                                                                                                                                                                                         (POTENTIAL)
                                                                                                                                                                                                                                                                                                                   (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                          (POTENTIAL)
                                                                                                                                                                                                                                         (POTENTIAL)
                                                                                                                                                                                                                                                                     (POTENTIAL)
                                                                                                                                                                                                                                                                                               (POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                        DB 1; Length 3010;
                                                                                                                                                                                                                                                                                                                                                                                                                                           11; Indels
                                                                                                                                                                                                                                                                                                                                                                                                      94A1C77435D642BB CRC64;
                                                                                                                                                     POTENTIAL.
CHARGE RELAY SYSTEM (BY
CHARGE RELAY SYSTEM (BY
CHARGE RELAY SYSTEM (BY
                                                                                                                                                                                                                                                                                                                                                                                                                       82.0%; Score 837; DB 1;
86.5%; Pred. No. 6.3e-71;
                                                                                                                                                                                                                                                                                                                                                                                  (GLCNAC.
                                                                                                                                                                                                                      (GLCNAC.
(GLCNAC.
(GLCNAC.
                                                                                                                                                                                                                                                         (GLCNAC.
                                                                                                                                                                                                                                                                                      (GLCNAC.
                                                                                                                                                                                                                                                                                                                                     (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                      (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                 (GLCNAC
                                                                                                                                                                                                             (GLCNAC
                                                                                                                                                                                                                                                  (GLCNAC
                                                                                                                                                                                                                                                                              (GLCNAC
                                                                                                                                                                                                                                                                                                         (GLCNAC
                                                                                                                                                                                                                                                                                                                           (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                          (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                             GLCNAC
                                                                                                                                                                                           ATP (POTENTIAL)
DECH BOX.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         PRT; 3033 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                           13; Mismatches
                                                                                                                                                                                                           N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                            N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                      326573
                                                                                                                                                                                                                                                                                                                                                                                                                                 al Similarity 86.5
154; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                       À,
                                                                                                                                                                                                                                                                                                                                                       2041
2077
2240
2529
2788
3010
                                                                                                                                                             1083
111083
111083
11168
1236
2250
2250
4417
4423
                                      3D-structure.
INIT_MET
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        POLG_HCVJ6
P26660;
                                                                                                                                                                                                                                                                                                                  CARBOHYD
                                                                                                                                                                                                                                                                                                                                     CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   78
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         138
                                                                                                                                                                                                                                                                                                                                                                                                                                                              18
                                                                                                                                                                                                                                                                                       CARBOHYD
                                                                                                                                                                                                                                                                                                                                                        CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                  CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                          CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                  CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                          Query Match
                                                                                                                                                                                         NP_BIND
SITE
                                                                                                                                                                                                                                                   CARBOHYD
                                                                                                                                                                                                                                                            CARBOHYD
                                                                                                                                                                                                                                                                     CARBOHYD
                                                                                                                                                                                                                                                                              CARBOHYD
                                                                                                                                                                                                                                                                                                 CARBOHYD
                                                                                                                                                                                                                                                                                                         CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                       SEQUENCE
                                                                                                                                                      TRANSMEM
                                                                                                                                                                         ACT_SITE
                                                                                                                                                                                                              CARBOHYD
                                                                                                                                                                                                                         CARBOHYD
                                                                                                                                                                                                                                  CARBOHYD
                                                                                                                                                                                                                                         CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                    Local
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      RESULT 7
POLG HCWJ6
TID POLG H
AC P2660
DT 01-AUG
DT 01-AUG
DT 02-FEB
DE Genome
DE ENVELO
DE (GP68)
DE (GP68)
DE (EC 3.
                                                                           CHAIN
CHAIN
CHAIN
CHAIN
CHAIN
                                                                                                                                    CHAIN
                                                                   CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                   Best Loc
Matches
  a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ŏ
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             MEDLINE-92044440; PubMed-1658196; Okamoch H., Okada K., Lizuka H., Okamoch H., Okada S.-I., Sugiyama Y., Kurai K., Lizuka H., Machida A., Miyakawa Y., Mayumi M.; Machida A., Miyakawa Y. Mayumi M.; "Nucleotide sequence of the genomic RNA of hepatitis C virus isolated
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Polyprotein; Glycoprotein; Transferase; RNA-directed RNA polymerase; Core protein; Coat protein; Envelope protein; Helicase; ATP-binding; Transmembrane; Nonstructural protein; Hydrolase; Serine proteise: INTL-MET
NS5B (P27); Nonstructural protein NS5A (P56); Nonstructural protein NS5B (P66) (P70) (RNA-directed RNA polymerase) (EC 2.7.7.48)]. Hepatitis C virus (isolate HC-J6) (HCV).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     -!- SUBDNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY LIPOPROTEIN BY THE ENVELOPE CONSIGNS OF TWO PROTEINS: PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MRNA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  .. SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY S29.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    | IPR001650; Helicase_C. | IPR001650; RNA_pol_DS_PS. | IPR007095; RNA_pol_DS_PS. | IPR007094; RNA_pol_DS_PS. | IPR007094; RNA_pol_DS_PS. | IPR007094; HCV_core; I. | IPR007099; HCV_NS1; I. | IPR007097; HCV_NS2; I. | IPR0097; HCV_NS3; I. | IPR0097; HCV_NS4a; I. | IPR0097; HCV_NS4b; I. | IPR0097; HCV_NS5a; I. | IPR0097; IIPR0097; 
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             HCV_capsid.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     HCV_NS4a.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              HCV_NS5a.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            HCV_core.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Pfam; PF00271; helicase_C; 1.
Pfam; PF00998; Viral_RdRP; 1.
ProDom; PD186062; HCV_NS1; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              EMBL; D00944; BAA00792.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SMART; SM00487; DEXDC;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            InterPro; IPR002518;
InterPro; IPR004109;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         InterPro; IPR000745;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             PIR; JQ1303; JQ1303.
HSSP; P27958; 1HEI.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              InterPro; IPR001410;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        InterPro; IPR002522;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           IPR002531;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    IPR001490;
                                                                                                                                                                                                                                                                                                                         SEQUENCE FROM N.A.
                                                                                                                                                                                                                                NCBI_TaxID-11113;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Pfam; PF01543;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              PF01539;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Pfam; PF01006;
Pfam; PF01001;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Pfam; PF01506;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       PF02907;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    039.00
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (RNA)(N).
                                                                                                                                                                                      Hepacivirus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              InterPro;
InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            InterPro;
```

```
NCBI_TaxID-11115;
                                                                                                                                                                                                                                                                                                                                                              (RNA)(N).
                                                                                                                                                                                    genotypes.
    1010 IGREVILGPADGYTSKGWSLLAPITAYAQQTRGLLGT1VVSMTGRDKTEOAGEIOVI.STV 1069
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            113 GSSDLYLVTRHADVIPVRRKGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVC 172
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              01-AUG-1992 (Rel. 23, Created)
01-AUG-1992 (Rel. 23, Last sequence update)
01-AUG-1992 (Rel. 23, Last sequence update)
01-AUG-1992 (Rel. 24). Last annotation update)
03-REB-2003 (Rel. 41, Last annotation update)
04-Operation (Rel. 41, Last annotation update)
05-Operation (Rel. 41, Last annotation update)
05-Operation (Rel. 41, Last annotation (P22);
05-Operation (Rel.);
05-Operation (Rel. 41, Nonstructural protein NS4 (Reparivirin)
05-Operation (Rel. 41, Rel. 41, Rel. 41, Nonstructural protein NS5 (PS5);
05-Operation (Rel. 41, Rel. 41, Rel. 41, Nonstructural protein NS5 (PS6);
05-Operation (Rel. 41, Rel. 41,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                10 VGRIVLNG -------AYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVSTA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    53 AQTFLATCINGVCWTVYHGAGTRTIASPKGPVIQMYTNVDKDLVGWPAPQGSRSLIPCTC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           17; Gaps
                                                                                                                    PROTEASE/HELICASE NS3 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4A (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4B (POTENTIAL).
RNA-DIRECTED RNA POLYMERASE (POTENTIAL).
                                   MATRIX PROTEIN (POIDWINIAL).
MAJOR ENVELOPE PROTEIN E (POTENTIAL).
NONSTRUCTURAL PROTEIN NS1 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS2 (POTENTIAL).
                                                                                                                                                                                                                   CHARGE RELAY SYSTEM (BY SIMILARITY)
CHARGE RELAY SYSTEM (BY SIMILARITY)
CHARGE RELAY SYSTEM (BY SIMILARITY)
                                                                                                                                                                                                                                                                                                                                                                                                                    (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                               (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Score 670.5; DB 1; Length 3033; Pred. No. 3.8e-55;
                    CAPSID PROTEIN C (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              F957F5C1A273BE9E CRC64;
  CELLULAR AMINOPEPTIDASE
                                                                                                                                                                                                                                                                                                                                     (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                              (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                    (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                           (GLCNAC.
                                                                                                                                                                                                                                                                                                                    (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              (GLCNAC
                                                                                                                                                                                                                                                                               ATP (POTENTIAL)
DECH BOX.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   PRT; 3033 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         27; Mismatches
                                                                                                                                                                                                    POTENTIAL
                                                                                                                                                                                                                                                                                                                    N-LINKED
N-LINKED
                                                                                                                                                                                                                                                                                                                                                                            N-LINKED
N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                    N-LINKED
N-LINKED
N-LINKED
N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                N-LINKED
N-LINKED
N-LINKED
N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    N-LINKED
N-LINKED
                                                                                                                                                                                                                                                                                                                                                              N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             173 TRGVAKAVDFIPVESLETIMRSP 195
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    65.7%;
62.1%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              329165
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   STANDARD;
                  2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

211

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2111

2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ź
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Query Match
Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    649
1091
2038
2811
3033
                                                                                                                                                                                                                                                                                                                                                                            305
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         126;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 POLG_HCVJ8
P26661;
                                                                                                                                                                                                                                                                                                                                                                                                                    CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CARBOHYD
                                                                                                                                                                                                                                                            ACT_SITE
NP_BIND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SEQUENCE
                                                                                                                                                                                                                                                                                                                    CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                              CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                   CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                           CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CARBOHYD
                                                                                                                                                                                                                                                                                                                                         CARBOHYD
                                                                                                                                                                                                                                                                                                                                                              CARBOHYD
                                                                                                                                                                                                    FRANSMEN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        RESULT 8
POLG_HCVJ8
                                                                                                                                        CHAIN
                                                                              CHAIN
                                                                                                                      CHAIN
                                                                                                                                                             CHAIN
                                                                                                                                                                                    CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
Son a serian a serian
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  d
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ð
```

```
the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way mondified and this statement is not removed. Usage by and for commercial entities requires a license agreement (Sce http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bloinformatics and the EMBL outstation -
                                                                                                                                                                                                                                                                                                                                                                                                                                            in the viral
                                                                              Okamoto H., Kurai K., Okada S.-I., Yamamoto K., Lizuka H., Tanaka T.,
Fukuda S., Tguda F., Mishiro S.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     -! - SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Core protein; Coat protein; Envelope protein; Helicase; ATP-binding;
Transmembrane; Nonstructural protein; Hydrolase; Serine protease.
INIT_MET 1 REMOVED FROM CAPSID PROTEIN C BY THE CELULAR AMINOPEPTIDASE.
CHAIN 116 191 MATRIX PROTEIN (POTENTIAL).
CHAIN 192 383 MAJOR ENVELOPE PROTEIN E (POTENTIAL).
CHAIN 384 733 NONSTRUCTURAL PROTEIN NSI (POTENTIAL).
CHAIN 734 1010 NONSTRUCTURAL PROTEIN NSI (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SWART; SM00487; DEXDc; 1.
Polyprotein; Glycoprotein; Transferase; RNA-directed RNA polymerase;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                    precursor polyprotein, commonly with Asp or Glu in the P6
position, Cys or Thr in P1 and Ser or Ala in P1'.
-!- CATALYTIC ACTIVITY: N nucleoside triphosphate - N diphosphate +
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS: PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MRNA.
                                                                                                                                                                *Full-length sequence of a hepatitis C virus genome having poor homology to reported isolates: comparative study of four distinct
                                                                                                                                                                                                                                                                       VICOLOGY 188:331-341(1992).

-!- FUNCTION: THE SMALL PROTEINS NS2A, NS2B, NS4A AND NS4B ARE HYDROPHOBIC, SUGGESTING A POSSIBLE MEMBRANE-RELATED FUNCTION.

NS3 AND NS5 MAY PLAY A ROLE IN THE VIRAL RNA REPLICATION.

-!- CATALYTIC ACTIVITY: Hydrolysis of four peptide bonds in the v
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 -! - SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY S29.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CO IPRO0252,
CO IPRO02519; HCV_env.
TO: IPRO02519; HCV_NS1.
PTO: IPRO02519; HCV_NS2.
PTO: IPRO04109; HCV_NS3.
CPTO: IPRO04109; HCV_NS4.
SIPPO: IPRO04409; HCV_NS4.
CEPTO: IPRO0456; HCV_NS4.
CEPTO: IPRO0746; HCV_NS5A.
CEPTO: IPRO07095; HCV_RAP.
CEPTO: IPRO07095; HCV_RAP.
CEPTO: IPRO07095; HCV_RAP.
CEPTO: IPRO07094; RNA_POL_PSVIT.
Fam; PFO1543; HCV_env: 1.
Pfam; PFO1543; HCV_env: 1.
Pfam; PFO1542; HCV_ENV: 1.
Pfam; PFO1560; HCV_NS1: 1.
Pfam; PFO1006; HCV_NS1: 1.
Pfam; PFO1001; HCV_NS3: 1.
Pfam; PFO1001; HCV_NS3: 1.
Pfam; PFO1001; HCV_NS4: 1.
SEQUENCE FROM N.A.
MEDLINE-92230232; PubMed=1314459;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  HCV_capsid.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                EMBL; D10988; BAA01761.1; -. PIR; A40250; GNWVJ8.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     InterPro; IPR001410; D:
InterPro; IPR002522; H:
InterPro; IPR002521; H:
InterPro; IPR002519; H
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          THEI.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              HSSP; P27958; 1HE
MEROPS; S29.001;
MEROPS; U39.001;
```

```
48984 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           STANDARD;
                                                                                                                                                                                                                                                                                                                                                                24
24
164
196
253
318
444
485 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                Complete proteome
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  NCBI_TaxID-3702;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         HHOA_ARATH
                                                                                                                                                                                                                                                                                                                                                                                                                                                                            SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                CHAIN
LIPID
DOMAIN
DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                             DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                            DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         RESULT 10
HHOA_ARATH
                                                                                                                                                                                                                                                                                                                                                   SIGNAL
   q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ô
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ä
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       53 AQTFLATCINGVCWTVYHGAGTRTIASPKGPVIQMYTNVDKDLVGWPAPQGSRSLTPCTC 112
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        25
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      10 VGRIVLNG -------AYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVSTA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    17; Gaps
                                           RNA-DIRECTED RNA POLYMERASE (POTENTIAL)
POTEWITAL.
CHARGE RELAY SYSTEM (BY SIMILARITY).
CHARGE RELAY SYSTEM (BY SIMILARITY).
ATP (POTEWITAL).
DECH BOX.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Fraser C.M., Norils S.T., Weinstock G.M., White O., Sutton G.G., Odoson R., Gwinn K.A., Sodorgon E., Gwinn K.A., Sodorgren E., Hardham J.M., McLood M.P., Salzberg S., Peterson J., Khalak H., Richardson D., Howell J.K., Chidambaram M., Utterback T.,
               NONSTRUCTURAL PROTEIN NS48 (POTENTIAL)
NONSTRUCTURAL PROTEIN NS4B (POIENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Treponema pallidum.
Bacteria: Spirochaetes: Spirochaetales: Spirochaetaceae; Treponema
NCBL_TaxID=160;
                                                                                                                                                                                                                                                 (POTENTIAL).
(POTENTIAL).
(POTENTIAL).
                                                                                                                                                                                                                                                                                                                               (POTENTIAL).
(POTENTIAL).
(POTENTIAL).
                                                                                                                                                                                                                                                                                                (POTENTIAL).
                                                                                                                                                                                                      (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                          (POTENTIAL)
 PROTEASE/HELICASE NS3 (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Length 3033;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                        INKED (GLCNAC. . .) (POTINKED (GLCNAC. . .) (POTINKED (ALCNAC. . .) (POTINAL)
                                                                                                                                                                                                                                    (GLCNAC. . . ) (GLCNAC. . . )
                                                                                                                                                                                                      (GLCNAC. . .)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      65.5%; Score 668.5; DB 1; 62.6%; Pred. No. 5.9e-55;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      25; Mismatches 34;
                                                                                                                                                                                                                                                                                                                                                                              (GLCNAC. .
                                                                                                                                                        (GLCNAC.
                                                                                                                                                                                                                                                                                                                                               (GLCNAC.
                                                                                                                                                                                                                                                                                                (GLCNAC.
                                                                                                                                                                                                                                                                                                                                  (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                             (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    16-OCT-2001 (Rel. 40, Created)
16-OCT-2001 (Rel. 40, Last sequence update)
16-OCT-2001 (Rel. 40, Last annotation update)
Hypothetical lipoprotein TP0136 precursor.
                                                                                                                                                                                                                                                                                                                                                                                                                             (GLCNAC
                                                                                                                                                        N-LINKED
N-L
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                173 TRGVAKAVDFIPVESLETTMRSP 195
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                STRAIN-Nichols;
MEDLINE-98332770; PubMed-9665876;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                          330177
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Matches 127; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                          1091
2038
2359
2811
                                                                                                                                                                                                                                                                                                                                                                                                                                                       2811 28
3033 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Y136_TREPA
083172;
                                                                                                                                                                                                                                                                                                                                CARBOHYD
CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                               CARBOHYD
                                                                                                         ACT_SITE
NP_BIND
                                                                                                                                                                                                                                                                                                                                                                                                             CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Query Match
                                                                                                                                                        CARBOHYD
                                                                                                                                                                                                       CARBOHYD
                                                                                                                                                                                                                        CARBOHYD
                                                                                                                                                                                                                                       CARBOHYD
                                                                                                                                                                                                                                                     CARBOHYD
                                                                                                                                                                                                                                                                     CARBOHYD
                                                                                                                                                                                                                                                                                     CARBOHYD
                                                                                                                                                                                                                                                                                                  CARBOHYD
                                                                                                                                                                                                                                                                                                                   CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                            CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                           CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                            SECUENCE
                                                                                                                                                                           CARBOHYD
                                                                                                                                                                                        CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Best Local
                                                              TRANSMEM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       rP0136.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             à
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its was by non-profit institutions as long as its content is in no way endified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     53 KAGSKLYATNGRL-----WEKELNGTGSWQKVSSSSVPTDSDK-----KVMSIATDGNTF 102
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 106 SLTPCTC----GSSDLYLVTRHADVIP-----VRRRGDSRGSLLSPRPISYLK--- 149
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            57 LATCI -- NGVCWTVYHGAG---TRIIASPKGPVIQMYINVDKDLVG------WPAPQGSR 105
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 159 NNGNCGCGGGGGGSSSSSSCIHIWLVPGGTGNNGNCGCGGGGGGSSSSSSCIHIKVEN 218
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     3 KRGSVVIV--GRIVLNGAYAQQTRGEEGCQETSQ----TGRDKNQVEGEVQIVSTAAQTF
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Lensch M.H.A., Sokolenko A., Herrmann R.G.;
"Identification and characterization of the an homolog to the bacterial periplasmic protease HoA.";
Submitted (DEC-1998) to the EMBL/GenBank/DDBJ databases.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Arabidopsis thallana (Mouse-ear cress).
Eukaryota, Viridiplantae, Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnollophyta; eudicotyledons; core eudicots; Rosidae;
eurosids II; Brassicales; Brasslcaceae; Arabidopsis.
McDonald L., Artiach P., Bowman C., Cotton M.D., Fujil C., Garland S.
Hatch B., Horst K., Roberts K., Sandusky M., Weldman J., Smith H.O.,
Venter J.C.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              71; Gaps
                                                                                                                  spirochete.';
Science 281:375-388(1998).
'' SUBCELLULAR LOCATION: Attached to the membrane by a lipid anchor
                                                                                      Complete genome sequence of Treponema pallidum, the syphilis
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               HYPOTHETICAL LIPOPROTEIN TP0136.
N-ACYL DIGLYCERIDE (POTENTIAL).
GLY/SER-RICH.
                                                                                                                                                                                                  (Potential).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           8.4%; Score 85.5; DB 1; Length 485; 3.6%; Pred. No. 1.3;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           78; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       C7A4CEEDC7DC5CED CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             -----GSSGGPLLCPAGHAVG 165
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              219 IDEQFLDMGEGYVVTTKHLYTKNGSSSAGPAQCPGGGGGG 258
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         EMBL; AE001199; AAC65137.1; ALT_INIT.
TIGR; TP0136; -.
Hypothetical protein; Lipoprotein; Membrane; Signal;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   09SEL7. 049507;
16-OCT-2001 (Rel. 40, Last sequence update)
16-OCT-2001 (Rel. 42, Last sequence update)
15-SEP-2003 (Rel. 42, Last annotation update)
Protease HhoA, chloroplast precursor (EC 3.4.21.-).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  321 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Query Match
Best Local Similarity 23.6%; Preu. ....
Matches 52; Conservative 19; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    GLY/SER-RICH.
POLY-SER.
POLY-SER.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                GLY/SER-RICH
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        POTENTIAL.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       a homolog 1
Submitted
```

```
RC STRAIN-C. Columbia.

RA MADELE 20083488; PubMed-10617198;
RA MEDICE-20083488; PubMed-10617198;
RA MADELE-20083488; PubMed-10617198;
RA MADELE-2008348; PubMed-10617198;
RA MADELE-2008349; PubMed-10617198;
RA MADELE-2008349; PubMed-10617198;
RA MADELE-2008349; PubMed-10617199; PubMed-10617199;
RA MADELE-2008349; PubMed-10617199; PubMed-10617199;
RA MADELE-2008349; PubMed-10617199; PubMed-10617199;
RA MADELE-2008349; PubMed-10617199; PubMed-10617199; PubMed-10617199; PubMed-1061799;
RA MADELE-2008349; PubMed-1061799; PubMed-106179; PubMe
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Nature 402:769-777(1999).
```

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its mee by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).

EMBL: AF114386; AAF24060.1; -. EMBL: AL021710; CAA16717.1; ALT\_SEQ. EMBL: AL161548; CAB78839.1; ALT\_SEQ.

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation-the European Bioinformatics Institute. There are no restrictions on its mose by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@lsb-sib.ch).
                                                                                                                                                                                                                                                                             73 GTRIIASPKGPVIQMYTNVDKDLVGWPAPQGSRSLTPCTCGSSDLYLVTRHADVIPVRRR 132
                                                                                                                                                                                                                                                                                               ---SLLSPRPISYLK-----GSSGGPLLCPAGHAVGIF 167
                                                                                                                                                                                                                                                                                                                                                              220 GNPYGYENTLIIGVVSGLGREIPSPNGKSISEAIQTDADINSGNSGGPLLDSYGHTIGV- 278
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Gaps
                                                                  Hydrolase; Serine protease; Chloroplast; Thylakoid; Transit peptide. TRANSII (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NCBL_TaxID=9606;
                                                                                                                                                                                                                                                     39;
                                                                                                                                     CHARGE RELAY SYSTEM (POTENTIAL).
CHARGE RELAY SYSTEM (POTENTIAL).
CHARGE RELAY SYSTEM (POTENTIAL).
R -> G (IN REF. 1).
68DB81E0BD27A7A7 CRC64;
                                                                                                                                                                                                                       8.1%; Score 82.5; DB 1; Length 321; 26.2%; Pred. No. 1.6; tive 22; Mismatches 49; Indels 39
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             15-JUL-1998 (Rel. 36, Last sequence update)
28-FEB-2003 (Rel. 41, Last annotation update)
Angio-associated migratory cell protein
                                                                                                               PROTEASE HHOA.
                                                                                                                                                                                                                                                                                                                                                                                                        279 NTATFTRKGSGMSSGVNFAIPIDTVVRTV 307
                                                                                                                                                                                                                                                                                                                                                                                         168 RAAVCTR---GVAKAVDF-IPVESLETTM 192
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SEQUENCE FROM N.A., AND PARTIAL SEQUENCE
MEROPS; SO1.279;
InterPro: IPRO01940; Protease2C.
InterPro: IPRO01254; Ser_protease_Try.
Pfam; PF00089; trypsin; 1.
PRINTS; PRO06894; PROTEASES2C.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                MEDI,INE=95262124; PubMed=7743515;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (Rel. 36, Created)
(Rel. 36, Last sequ
                                                                                                                                                                                               34691 NW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              EMBL; M95627; AAA68889.1; -. PIR; I39383; I39383.
                                                                                                                                                                                                                                                     Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       STANDARD;
                                                                                                                                                                                                                                                                                                                                    133 GDSRG-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Genew; HGNC:18; AAMP.
MIM; 603488; -.
                                                                                                                                                                                               321 AA;
                                                                                                                                                                                                                                      Best Local Similarity
Matches 39; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      TISSUE-Brain;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              15-JUL-1998
15-JUL-1998
28-FEB-2003
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    AAMP_HUMAN
013685;
                                                                                                                         DOMAIN
ACT_SITE
ACT_SITE
ACT_SITE
CONFLICT
                                                                                                                                                                                             SEQUENCE
                                                                                                                                                                                                                         Query Match
                                                                                 TRANSIT
TRANSIT
CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                          AAMP_HUMAN
                                                                                                                                                                                                                                                                                                                                                                                                                                                             RESULT 11
  g
                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                              셤
                                                                                                                                                                                                                                                                                                                                                                                                                  g
```

```
MEMBRANE.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Hydrolase; S
TRANSIT
TRANSIT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ACT_SITE
ACT_SITE
CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ACT_SITE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  OMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        IBPI_HUMAN
ID IBPI_H(
AC P08833
DT 01-NOV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Matches
        a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        οy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            a
                                                                                                                                                                                                                                                                                                                                                                                                                                               6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     66 WTVYHGAGTRTIASPKGPV1QMYTNVDKDLVGWPAPQGSRSL-----TPCTCGSSDLYLV 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              121 TRHADVIPVRRR---GDSRGS----LLSPRPISYLKGSSG--GPLLCPA------160
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         241 -----VLPDGKRAVVGYEDGTIRIWDLKQGSP1HVLKGTEGHQGPLTCVAANQDGSL1LT 295
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               DEGI_ARAIH STANDARD: PRT; 437 AA.
022609; 09LK85;
16-0CT-2001 (Rel. 40, Created)
16-0CT-2001 (Rel. 40, Last sequence update)
28-FBB-2003 (Rel. 41, Last annotation update)
Protease Do-like 1, chloroplast precursor (EC 3.4.21.-),
PEGPI OR DEGP OR AT3627925 OR K16N12.18
Arabidopsis thaliana (Mouse-ear cress).
Bukaryota, Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; cudicotyledons; core eudicots; Rosidae;
eurosids II; Brassicales; Brassicaceae; Arabidopsis.
                                                                                                                                                                                                                                                                                                                                                                                                                                                 Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     STRAIN-CV COlumbia;

MEDLINE-20363099; PubMed-10907853;

Raneko T., Katoh T., Sato S., Nakamura A., Asamizu E., Tabata S.;

Kaneko T., Katoh T., Sato S., Nakamura A., Asamizu E., Tabata S.;

"Structural analysis of Arabidopsis thaliana chromosome 3. II.

Sequence features of the 4,251,695 bp regions covered by 90 Pl. TAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        STRAIN-cv. Columbia:
Kaselbach T. Bystedt M., Schroeder W.P.;
Submitted (JUL-2000) to the SWISS-PROT data bank.
-!- FUNCTION: SERINE PROTEASE THAT IS REQUIRED AT HIGH TEMPERATURE.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SEQUENCE FROM N.A., AND CHARACTERIZATION.
MEDLINE-98175982; PubMed-9507020;
Itzhaki H., Naveh I., Lindahl M., Cook M., Adam Z.;
Ildentification and characterization of DegP, a serine protease associated with the luminal side of the thylakoid membrane.";
J. Biol. Chem. 273:7094-7098(1998).
                                                                                                                                                                                                                                                                                                                                                                                                                                                 94;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    296 GSVDCQAKLVSATIGKVVGVFRPETVASQPSLGEGEESESNSVESL 341
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               161 ------GHAVGIFR----AAVCTRGVAKAVDFIPVESL 188
                                                                                                                                                                                                                                                                                                                                                                                                     8.0%; Score 82; DB 1; Length 452; 25.3%; Pred. No. 2.6;
                                                                                                                                                                                                                                                                                                                                                                                                                                             13; Mismatches 47; Indels
                                                                                                                                                                  HEPARIN-BINDING (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                 DA1413D25EB236C0 CRC64;
GO; GO:0008201; F:heparin binding activity; TAS.
InterPro; IPR001680; WD40.
Pfam; PF00400; WD40; 8.
SMART; SM00310; WD40; 8.
PROSITE; PS00678; WD_REPEATS_1; 1.
PROSITE; PS50082; WD_REPEATS_2; 6.
PROSITE; PS50294; WD_REPEATS_REGION; 1.
                                                                                                                                                                                       POLY-GLU.
                                                                                                                                                                                                       WD 2. WD 2. WD 4. WD 5. WD 6. 
                                                                                                                                                                                                                                                                                                                                                                 49015 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                          Best_Local Similarity 25.3%
Matches 42; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             DNA Res. 7:217-221(2000)
                                                                                                                                                                  18
77
138
180
220
220
261
363
404
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SEQUENCE OF 104-118.
                                                                                                                                                                                                                                                                                                                                                                 452 AA;
                                                                                                                                               Repeat; WD repeat.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          NCBI_TaxID=3702;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            BAC clones.
                                                                                                                                                                                                                                                                                                                                                                   SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                         Query Match
                                                                                                                                                                                     DOMAIN
REPEAT
REPEAT
REPEAT
                                                                                                                                                                                                                                                                REPEAT
REPEAT
                                                                                                                                                                                                                                                                                                                         REPEAT
REPEAT
                                                                                                                                                                    DOMAIN
                                                                                                                                                                                                                                                                                                          REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           RESULT 12
DEGI_ARATH
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            and
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ď
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               à
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMEL outstation the Buropean Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb.sib.ch)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                7;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 68 VYHGAGTRIIASPKGPVIQMY-----PA 100
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                101 PQGSRSLTPCTCGSSDLYLV-----TRHADVIPVRRRGDSRGSLLSPRPI 145
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      150 VPQGSGSGFVWDKQGHIVTNYHVIRGASDLRVTLADQTTFDAKVVGFDQDKDVAVLRIDA 209
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     210 PK--NKLRPIPVGVSADLLVGQKVFAIGNPFGLDHTLTTGVISGLRREIS--SAATGRPI 265
                                          CAN DEGRADE BETA-CASEIN.
ENZYME REGULATION: INHIBITED BY PHENYLMETHYLSULFONYL FLUORIDE AND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   PDZ
CHARGE RELAY SYSTEM (POTENTIAL).
CHARGE RELAY SYSTEM (POTENTIAL).
CHARGE RELAY SYSTEM (POTENTIAL).
HSPPSSQLSNST -> SSTLFLHSPPSSHL (IN REF
BE INVOLVED IN THE DEGRADATION OF DAMAGED PROTEINS. IN VIVO,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Transit peptide; Chloroplast; Thylakoid.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    146 SYL-----KGSSGGPLLCPAGHAVGIFRAAVCTRGVAKAVDF-IPVESL 188
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    O-PHENANTHROLINE.
-1- SUBCELLULAR LOCATION: BOUND TO LUMINAL SIDE OF THE THYLAKOID
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       55;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            7.9%; Score 80.5; DB 1; Length 437; 25.6%; Pred. No. 3.5; tive 18; Mismatches 55; Indels 59
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         2).
P -> I (IN REF. 2).
P -> S (IN REF. 2).
G -> R (IN REF. 2).
LL -> D (IN REF. 2).
LL -> V (IN REF. 2).
L -> V (IN REF. 2).
I -> V (IN REF. 2).

                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CHLOROPLAST (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              PROTEASE DO-LIKE 1. SERINE PROTEASE.
                                                                                                                                                                                                                                                                    -!- INDUCTION: By heat shock.
-!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S2C.
-!- SIMILARITY: Contains 1 PD2/DHR domain.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Ş
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                MEROPS; SO1.779; --
InterPro; IPR001478; PD2.
InterPro; IPR00140; Protease2C.
InterPro; IPR001254; Ser_protease_Try.
Pfam; PP00595; PD2, 1.
Pfam; PF00089; trypsin; 1.
Pfam; PR00189; trypsin; 1.
PRINTS; PR0084; PROTEASES2C.
SMART; SM00128; PD2; 1.
PROSITE; PS50106; PD2; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    EMBL; AF028842; AAC39436.1; -.
EMBL; AP000371; BAB02539.1; -.
EMBL; AP001302; BAB02539.1; JOINED.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               IBP1_HUMAN STANDARD; P08833; 01-NOV-1988 (Rel. 09, Created)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            46213 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   PS50106; PDZ; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       44; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    321
421
171
201
280
23
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  36
54
60
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             36
54
66
66
355
335
381
416
437 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Query Match
Best Local Similarity
```

```
DISULFIDE BONDS.
                                                                                                                                                                                                                                                                                                             MUTAGENESIS
   "Insulin-like growth factor (IGF) binding protein complementary deoxyribonucleic acid from human HEP G2 hepatoma cells: predicted protein sequence suggests an IGF binding domain different from those of the IGF-I and IGF-II receptors.";
                                                                                                                       SEQUENCE FROM N.A.
TISSUE-placenta;
THEDLINE-8905264: PubMed-2461284;
Brinkman A., Groffen C., Kortleve D.J., Geurts A., Drop S.L.S.;
Brinkman A., Groffen C., Kortleve D.J., Geurts A., Drop S.L.S.;
"Isolation and characterization of a cDNA encoding the low molecular weight insulin-like growth factor binding protein (IBP-1).";
EMBO J. 7:2417-2423(1988).
                                                                                                                                                                                                                                                                                                                                                                   Grundmann U., Nerlich C., Bohn H., Rein T.;
"Cloning of CDNA encoding human placental protein 12 (PP12): binding protein for IGF 1 and somatomedin.";
Nucleic Acids Res. 16:8711-8711(1988).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Julkunen M., Koistinen R., Aalto-Setala K., Seppala M., Janne O.A.,
                                                                                                                                                                                                                                                        Brewer M.T., Stetler G.L., Squires C.H., Thompson R.C., Busby W.H. Jr., Clemmons D.R.; Cloning, characterization, and expression of a human insulin-like growth factor binding protein. 152:1289-1297(1988).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Kontula K.;
"Primary structure of human insulin-like growth factor-binding
protein/placental protein 12 and tissue-specific expression of its
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SEQUENCE FROM N.A. MEDLINE-89087480; PubMed-2849945; MEDLINE-89087480; PubMed-2849945; Medline-8., Gorfen C.A., Kortleve D.J., Drop S.L.; "Organization of the gene encoding the insulin-like growth factor binding protein IBP-1.*; Biochem. Blophys. Res. Commun. 157:898-907(1988).
                                                                   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NCBI_TaxID=9606;
            28-FEB-2003 (Rel. 41, Last annotation update)
Insulin-like growth factor binding protein 1 precursor (IGEBP-1)
(IBP-1) (IGF-binding protein 1) (Placontal protein 12) (PP12).
IGEB1 OR IBP1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   [6]
SEQUENCE FROM N.A.
HEDLINE-89330502; PubMed-2474129;
Cubbage M.L., Suwanichkul A., Powell D.R.;
"Structure of the human chromosomal gene for the 25 kilodalton insulan-like growth factor binding protein.";
Mol. Endocrinol, 3:846-851(1989).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Lee Y. L., Hintz R.L., James P.M., Lee P.D.K., Shively J.E., Powell D.R.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Ehrenborg E., Larsson C., Stern I., Janson M., Powell D.R.,
 01-NOV-1988 (Rel. 09, Last sequence update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             of the IGF-I and IGF-II receptors."; Mol. Endocrinol. 2:404-411(1988).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                MEDLINE-92217971; PubMed-1373120;
                                                                                                                                                                                                                                                                                                                                                      MEDLINE-88335621; PubMed-3419931;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             [5]
SEQUENCE FROM N.A.
MEDLINE-88334540; Pubmed-2458522;
                                                                                                                                                                                                                                                                                                                                                                                                                                                          MEDLINE-88312985; PubMed-2457513;
                                                                                                                                                                                                                                                 MEDLINE-88240345; Pubmed-2454104;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  FEBS Lett. 236:295-302(1988).
                                                                Homo sapiens (Human).
                                                                                                                                                                                                                         SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                   SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                  SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                TISSUE-Placenta;
                                                                                                                                                                                                                                      TISSUE-Decidua;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        HRNA
```

```
Neumann G.M., Bach L.A.;
Neumann G.M., Bach L.A.;
The N-terminal disulfde linkages of human insulin-like growth
Neumann G.M., Bach L.A.;
The N-terminal disulfde linkages of human insulin-like growth
Transport of the disulfde linkages of human insulin-like growth
Transport of the disulfde linkages of human insulin-like growth
Transport of the display of the different as and have been shown to either inhibit or stimulate the growth
Transport of IGF binding proteins prolong the half-life of the IGFs and have been shown to either inhibit or stimulate the growth promoting effects of the IGFs on Cell culture. They alter the interaction of IGFs with their cell surface receptors.

1. SIMCELLULAR LOCATION: Secreted.
They alter the affinity of the protein for IGFs.
They is the protein for IGFs.
They contains 1 IGFs domain.
They contains 1 IGFs domain.
They alter the frameshifts in positions 55 and 71.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            This SWISS-PROT entry is copyright. It is produced through a collaboration between the SWISS Institute of Bloinformatics and the EMBL outstation the Buropean Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                   MEDINE-8017023; PubMed-2466665;
Luthman H., Soederling-Barros J., Persson B., Engberg C., Stern I.,
Luthman H., Soederling-Barros J., Persson B., Endgren B.,
Hjelmqvist L., Enerbaeck S., Carlsson P., Bjursell G., Povoa G.,
Hjelmqvist L., Gerrorall H.,
"Human insulin-like growth-factor-binding protein. Low-molecular-mass form: protein sequence and cDNA cloning.";
[10]
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Increase of beta-actin mRNA upon hypotonic perfusion of perfused rat
*Contiguous localization of the genes encoding human insulin-like growth factor binding proteins 1 (IGBP1) and 3 (IGBP3) on chromosome
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Busby W.H. Jr., Klapper D.G., Clenmons D.R.; "Purification of a 31,000-dalton insulin-like growth factor binding protein from human amnlotic fluid. Isolation of two forms with different biologic actions.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       "Identification of the sites of phosphorylation in insulin-like growth factor binding protein-1. Regulation of its affinity by phosphorylation of serine 101.";
J. Biol. Chem. 268:1125-1131(1993).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    MEDLINE-93123224; PubMed-7678248;
Jones J.I., Busby W.H. Jr., Wright G., Smith C.E., Kimack N.M.,
Clemmons D.R.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  MEDLINE-92070504; PubMed-1959616;
Brinkman A., Kortlrve D.J., Schuller A.G.P., Zwarthoff E.C.,
Drop S.L.S.;
                                                                                                                                                                                                                      SEQUENCE OF 141-259 FROM N.A., AND SEQUENCE OF 26-259 TISSUE-Amniotic fluid;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               PHOSPHORYLATION SITES, AND PARTIAL SEQUENCE.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Biol. Chem. 263:14203-14210(1988).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          EMBL, Y00856; CAA68770.1; -.
EMBL, M20841; AAA52540.1; ALT_FRAME.
EMBL; X12385; CAA30942.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       MEDLINE-99262603; PubMed-10329650;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SEQUENCE OF 26-53.
MEDLINE-89008261; PubMed-2971653;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  FEBS Lett. 292:264-268(1991).
                                                                                                                             Genomics 12:497-502(1992).
```

```
MEDLINE-86200439; PubMed-3009890;
                                                                                                                                                                                                                                                                                                          EMBL; AE004818; AAG07406.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                Ouery Match
Best Local Similarity 26.58
Matches 50; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           182 FIPVESLET 190
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    181 FVVARILNT 189
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ENV_FLVGL
P08359;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ENV_FLVGL
AC DE ENV_F
AC DI - AU
DI 10 - AU
DI 10 - AU
DI 10 - AU
DE ENV F
DE ENV F
GN ENV COX VICUS
OX VICUS
OX VICUS
OX NCBI
RN FELLI
RN (11)
RN SEQU
RX MEDL
RA STEW
RA STEW
RT "NUC
RT GENE
           q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             QQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ó
                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             12 RIVLNGAYAQQTRGEE-----GCQEIS--OIGRDKNQVEGEVQIVSTAAQTFLATCIN 62
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   28; Gaps
                                                                                                                                                                                                MIN: 146730; -. 60 of the space; TAS. 60 of GO:0005615; C:extracellular space; TAS. 60 of GO:0005520; F:insulin-like growth factor binding activity; TAS. 60; GO:0007165; P:signal transduction; TAS. InterPro; IPR000867; Insl_gro_fac_pr. InterPro; IPR000716; Thyrogiobulin_l.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
                                                                                                                                                                                                                                                                                                                                                                                         INSULIN-LIKE GROWTH FACTOR BINDING
                                                                                                                                                                                                                                                                                                                                 PROSITE; PS00222; IGF_BINDING; 1.
PROSITE; PS00484; THYROGLOBULIN_1; 1.
Growth factor binding: Signal; Phosphorylation; Polymorphism.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        7.7%; Score 79; DB 1; Length 259; 28.3%; Pred. No. 2.6;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      182 RVVESLAKAQETSGEEISKFYLPNCNKNGFYHSRQCETSMDGEA----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   12; Mismatches 26; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   PHOSPHORYLATION (MAJOR).
PHOSPHORYLATION.
V -> I (IN dbSNP:1065782).
//TTd-VAR. (IN dbSNP:4619).
/FTId-VAR_003821.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              -> Q (IN REF. 2).
8AA75AF7DC966012 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                    PROTEIN 1.
THYROGLOBULIN IYPE I.
CELL ATTACHMENT SITE.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     16-OCT-2001 (Rel. 40, Created)
16-OCT-2001 (Rel. 40, Last sequence update)
16-OCT-2001 (Rel. 40, Last annotation update)
Probable aromatic acid decarboxylase (RC 4.1.1.-).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                63 GVCWTVYHGAGTRIIASPK---GPVIOMYTNV 91
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                209 AA
PRT;
                                                                                                                                                                                                                                                                            Pfam: PF00219; IGFBP; 1.
Pfam: PF00086; Uhyroqlobulin_1: 1.
SMART: SW00121; IB: 1.
SMART: SW00211; TY: 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Pseudomonadaceae; Pseudomonas
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              27903 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Ouery Match
Best Local Similarity 28.3%
Matches 26; Conservative
                                                                                                                                                                Siena-2DPAGE; P08833; -...
Genew; HGNC:5469; IGFBP1.
MIM; 146730; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Pseudomonas aeruginosa
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                213 2
259 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SEQUENCE FROM N.A.
                                                                                                                                                          FRANSFAC; T00400;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       253
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      NCBI_TaxID=287;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             PAAD_PSEAE
Q9HX08;
                                                                                                                                                                                                                                                                                                                                                                                                                                           DISULFID
DISULFID
DISULFID
MOD_RES
MOD_RES
MOD_RES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       VARIANT
                                                                                                                                                                                                                                                                                                                                                                                                                   DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                PA4019
                                                                                                                                                                                                                                                                                                                                                                                         CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             PAAD_PSEAE
```

a

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMEL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               84 VIQMYTNVDKDLVGWPAPQGSRSLTP-----CTCGSSDL-----YLVTRHADVIPV 129
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         75 QIRVFGQND-----WMAPPASGSSAPNAMVICPCSTGTLSAVATGACNNLIERAADVALK 129
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           83
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   29 QEEREVHFLISKAAQLVMAT------ETDVALPAKPQAMQAFLTEYCGAAAG 74
STRAIN-ATCC 15692 / PAO1;
MEDLINE-20437337; PubMed-10984043;
Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warrener P.,
Eickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,
Garber R.L., Goltry L., Tolentino E., Westbrock-Wadman S., Yuan Y.,
Brody L.L., Coulter S.N., Folger K.R., Kas A., Larbig K., Lim R.M.,
Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
Reizer J., Saler M.H., Hancock R.E.W., Lory S., Olson M.V.;
"Complete genome sequence of pseudomonas aeruginosa PAO1, an
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              130 RRRGDSRGSLLSPR--PIS----YLKGSSGGPLLCPAGHAVGIFRAAVCTRGVAKAVD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       41 QVEGEVQ-IVSTAAQIFLATCINGVCWIVYHGAGIRIIASPKGP---------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Onions D.E., Neil J.C.;
"Nucleotide sequences of a feline leukemia virus subgroup A envelope gene and long terminal repeat and evidence for the recombinational origin of subgroup B viruses.";
J. Virol. 58:825-834(1986).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      67;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         PIR: H83144; H83144.
InterPro; IPR003382; Flavoprotein.
Pfam, PF02441; Flavoprotein; 1.
Hypothetical protein; Lyses; Decarboxylase; Complete proteome.
SEQUENCE 209 AA; 22367 MW; 01FD081CC495D3F6 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Stewart M.A., Warnock M., Wheeler A., Wilkie N., Mullins J.I.,
                                                                                                                                                                                                                                                                                                                                                                                                                                         Nature 406:959-964(2000).
-:- SIMILARITY: BELONGS TO THE POLYPRENYL P-HYDROXYBENZOATE / PHENYLACRYLIC ACID DECARBOXYLASES FAMILY.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          01-AUG-1988 (Rel. 08, Created)
01-AUG-1988 (Rel. 08, Last sequence update)
16-CCT-2001 (Rel. 40, Last annotation update)
ENV polyprotein precursor (Coat polyprotein) (Contains: Knob (GP70; Spike protein P15E).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        7.7%; Score 78.5; DB 1; Length 209; 26.5%; Pred. No. 2.3; tive 16; Mismatches 56; Indels 6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Feline leukemia virus (strain A/Glasgow-1).
Viruses; Retroid viruses; Retroviridae; Gammaretrovirus.
NCBI_TaxID=11769;
```

```
between the Swiss institute of Bioinformatics and the FMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
This SWISS-PROT entry is copyright. It is produced through a collaboration
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 æ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             326 IAILGNYSNQTNPPPSCLSTPQHKLTISEVSGGGMCIGTVPKTHQALCNK----TQGGHT 383
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            73 GTRTIASPKGPVIQMYTNVDKDLVGWPAPQGSRSL/FPCT------CGSSDL,--- 117
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            -----YLVTRHADVIPVARRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAV 171
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               IVLNGAYAQQIRGEEGCQETSQTGRDKNQVEGEVQIVSTAAQIFLATCINGVCWTVYHGA 72
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Indels 59; Gaps
                                                                                                                                                                                                                      KNOB PROTEIN GP70.

SPIKE PROTEIN P15E.

N-LINKED (GLCNAC. .) (POTENTIAL).

N-LINKED (GLCNAC. .) (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  DB 1; Length 642;
                                                                                                                       EMBL: M12500: AAA43053.1: -
HSSP; P03385: 1MOF.
InterPro; IPR002059: Env_polyprotein.
Fdam; PF00429: ENV_polyprotein; 1.
Coat protein: Glycoprotein; Polyprotein; Signal.
SIGNAL
34 445 KNOB PROTEIN GP70.
CHAIN 446 642 SPIKE PROTEIN GP70.
CARBOHYD 35 35 N-LINKED (GLCNAC...)
CARBOHYD 43 43 N-LINKED (GLCNAC...)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               16; Mismatches 83;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                7.5%; Score 77.5; 1
22.2%; Pred. No. 10;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Search completed: August 30, 2603, 19:13:44
Job time : 10.6577 secs
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    GT-GTKALLETAQFROLOMAMHT 490
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CTRGVAKAVDFIPVESLETTMRS 194
                                                                                                                                                                                                                                                                                                                                                                                                                                  71053 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Conservative
                                                                                                                                                                                                                           410
642 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Ma.
Local Sim
45;
                                                                                                                                                                                                                                                                                                     CARBOHYD
                                                                                                                                                                                                                                                                                                                                     CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 13
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            382
                                                                                                                                                                                                                                                                                                                                                                                                                   CARBOHYD
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          118
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Query Match
                                                                                                                                                                                                                                                                                        CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                      CARBOHYD
 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ΔQ
```

hepatitis hepatitis

Q91RS3 Q91RT4 Q91RS8 Q91RT3 Q91RS7 Q91RS7 Q91RT0 Q90TT6 Q90TT6

hepatitis hepatitis hepatitis hepatitis hepatitis hepatitis hepatitis hepatitis hepatitis

hepatitis hepatitis hepatitis hepatitis hepatitis hepatitis hepatitis hepatitis hepatitis

091rrs 09

Q9PWU9 Q91RX4 Q91RTF Q91RTP Q91RRQ Q91RRQ Q91RRQ Q91RRQ Q91RRQ Q91RRQ Q91RRD Q91RRD Q91RRD Q91RRD Q91RRD Q91RRD

hepatitis hepatitis hepatitis hepatitis hepatitis

Q91RT7 Q91RS0 Q91RT8 Q81754

```
SEQUENCE FROM N.A.
MEDLINE-2013484; PubMed-10864644;
Lai V.C., Zhong W., Skelton A., Ingravallo P., Vassilev V.,
Donis R.O., Hong Z., Lau J.Y.;
"Generation and characterization of a hepatitis C virus NS3 protease-
dependent bovine viral diarrhea virus.";
J. Virol. 74:6339-6347(2000).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              A Lai V.C.H., Hong 2.;

Lai V.C.H., Hong 2.;

Submitted (MAY-2000) to the EMBL/GenBank/DDBJ databases.

R RESPERT & ARF82566.1;

R HSSP: P66663: JJXP.

R INTEPPO: JRR000280. CDvir_endptseP80.

R InterPro: JRR000280. CDvir_endptseP80.

R InterPro: JRR001109. DEAD.

R InterPro: JRR001109. DEAD.

R InterPro: JRR001109. Helicase_C.

R InterPro: JRR001109; Helicase_C.

R InterPro: JRR001055; WAB_PDNA_binding.

R InterPro: JRR001055; RNA_pol_DS_PS.

R InterPro: JRR001055; RNA_pol_DS_PS.

R InterPro: JRR001095; RNA_pol_DS_PS.

R InterPro: JRR001094; RNA_pol_DS_PS.

R Pfam: PF002907; HCV_NS3: 1.

Pfam: PF002907; HCV_NS3: 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Last sequence update)
Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Viruses; ssRNA positive-strand viruses, no Pestivirus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Created)
 01-OCT-2000 (TrEMBLrel.
01-OCT-2000 (TrEMBLrel.
01-MAR-2003 (TrEMBLrel.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         PRELIMINARY;
Genome polyprotein.
Mucosal disease virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   [2]
SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             NCBI_TaxID=11099;
 876.5
876.5
876
Q9IFH8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Q91FH8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     RESULT 1
091FH8
 DORRERTH SERVICE DESTRICT DEST
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  nucosal dis
hepatitis c
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           0 0 0 0 0 0 0 0 0 0 0 0 0 0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               hepatitis of hepat
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               hepatitis
                                                                                                                                      (without alignments)
1352.314 Million cell updates/sec
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.
                                                                                                                                                                                                                        ......VAKAVDFIPVESLETTMRSP 195
                                                                                                                    August 30, 2003, 19:00:22; Search time 37:2105 Seconds
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Description
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Q91fe5
Q03463
Q031rr8
Q91rr5
Q91rr3
Q91rr3
Q91rr3
Q91rr6
Q91rr6
Q91rr6
Q91rr6
Q91rr6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     09ifh8
Q81756
                                                                                                                                                                                                                                                                                                                                              830525
                GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.
                                                                                                                                                                                                                                                                                                                                            Potal number of hits satisfying chosen parameters:
                                                                                                                                                                                                                                                                                                              830525 seqs, 258052604 residues
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SUMMARIES
                                                                                                                                                                                                                                                                                                                                                                                                                              Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries
                                                                                       OM protein – protein search, using sw model
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     091FH8
081756
003465
0091RR8
091RR7
091RR7
091RR3
091RR1
091RR1
091RR6
091RR1
091RR6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     sp_plant:*
sp_rodent:*
sp_virus:*
sp_vertebrate:*
sp_unclassified:*
sp_rvirus:*
                                                                                                                                                                                                                                                          BLOSUM62
Gapop 10.0 , Gapext 0.5
                                                                                                                                                                                                                        1 MKKKGSVVIVGRIVLNGAYA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               sp_bacteria:*
sp_fungi:*
sp_human:*
sp_huwatebrate:*
sp_mammal:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 sp_bacteriap:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         sp_mhc:*
sp_organelle:*
                                                                                                                                                                                                                                                                                                                                                                                length: 0
length: 2000000000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 sp_archeap:
                                                                                                                                                                                 US-09-965-594-12
1021
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       DB
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    sp_archea:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SPTREMBL_23:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            sp_phage:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Query
Match 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     10:
                                                                                                                                                                                                                                                                                                                                                                                                                                  Post-processing:
                                                                                                                                                                                        Title:
Perfect score:
Sequence:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Minimum DB seq
Maximum DB seq
                                                                                                                                                                                                                                                            Scoring table:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Score
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ..
                                                                                                                                                                                                                                                                                                              Searched:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Database
                                                                                                                       Run on:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Result
```

PRT; 4040 AA

ALIGNMENTS

DNA stage; Flaviviridae;

~

```
InterPro;
InterPro;
InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  InterPro;
InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        InterPro;
                                                    NON_TER
SEQUENCE
                                                                                    Query Match
                                                                                                                                                                                                                                                                                                                                        Q91FE5
                                                                                                                                                                                                                                                                                                                 RESULT 3
Q9IFE5
  DR W W W T L C O
                                                                                                                               ò
                                                                                                                                                  qq
                                                                                                                                                                         ò
                                                                                                                                                                                             q
                                                                                                                                                                                                                  õ
                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                 qq
                                                                                                                                                                                                                                                                                                                                                 CINGVCWTVYHGAGIRTIASPKGPVIQMYTNVDKDLVGWPAPQGSRSLTPCTCGSSDLYL 119
                                                                                                                                                                                                                                                 120 VTRHADVIPVRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVCTRGVAKA 179
                                                                                                                                                              29
                                                                                                                                                                         10 GSVVIVGRIVLSGSGSITACAQQTRGLLGCKITSLTGRDKNQVEGEVQIVSTATQTFLAT 69
        SWART: SM00487; DEXDC; 1.

SMART: SM00490; HELLCC; 1.

PROSITE; PS00037; WNR=1: 1.

PROSITE; PS50521; RDRP_VIRAL; 1.

PROSITE; PS50521; RDRP_VIRAL; 1.

ATP-binding; Hellcase: Hydrolase; Norstructural protein; Polyprotein; RNA-directed RNA polymerase; Transferase.

SEQUENCE 4040 AA; 453073 WW; ADE87791D05589DC CRC64;
                                                                                                                                                              5 GSVVIVGRIVLNG----AYAQQTRGEEGCQETSQTGRDKNCVEGEVQIVSTAAQTF1,AT
                                                                                                                                         Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                          Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
Hepacivirus.
                                                                                                                                          ..
                                                                                                                  89.0%; Score 908.5; DB 12; Length 4040; 90.8%; Pred. No. 9e-84; ative 6; Mismatches 7; Indels 5;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Choo Q.-L., Richman K., Han J.;

"The nucleotide sequence of the Hepatitis C viral genome.";

"The nucleotide sequence of the EMBL/GenBank/DDBJ databases.

EMBL: M32084; AA45677.1; -.

HSSP: P27958; IAIV.
                                                                                                                                                                                                                                                                                                                                                                                        01-NOV-1996 (TrEMBLrel. 01, Created)
01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
Genome polyprotein (Fragment).
Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                      PRT; 2436 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    INTERPOS TRANSTANTING USAD.
INTERPOS TPRO01410; DEAD.
INTERPOS TPRO01531; HCV_MS1.
INTERPOS TPRO04109; HCV_MS2.
INTERPOS TPRO04109; HCV_MS4.
INTERPOS TPRO01409; HCV_MS4.
INTERPOS TPRO01409; HCV_MS4.
INTERPOS TPRO01868; HCV_MS4.
INTERPOS TPRO0166; HCV_MS6.
INTERPOS TPRO0166; HCV_MS6.
INTERPOS TPRO01095; RNA_DO1_DS_FS.
INTERPOS TPRO07094; RNA_DO1_DS_FS.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Pfam; PPO1560; HCV_NS1: 1.
Pfam; PPO1578; HCV_NS2: 1.
Pfam; PPO1006; HCV_NS4: 1.
Pfam; PPO1006; HCV_NS4b; 1.
Pfam; PPO1001; HCV_NS4b; 1.
Pfam; PPO1506; HCV_NS5b; 1.
Pfam; PPO0921; holicase_C; 1.
Pfam; PPO0988; Viral_RGRP; 1.
PROSITE; SMO487; DEXDC; 1.
PROSITE; PS50507; RDRP_POSITIVE; 1.
PRINTS; PR00729; CDVENDOPTASE
                                                                                                                                                                                                                                                                                                        180 VDFIPVESLETTMRS 194
                                                                                                                              Best Local Similarity 90.8
Matches 177; Conservative
                                                                                                                                                                                                                                                                                                                                                                       PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                    NCBI_TaxID=11103;
                                                                                                                    Query Match
                                                                                                                                                                                                                                                                                                                                                                    Q81756
Q81756;
                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                               a
                                                                                                                                                                                                       ò
                                                                                                                                                                                                                      a
                                                                                                                                                                                                                                                                      qq
                                                                                                                                                                                                                                                                                            ö
                                                                                                                                                                                                                                                 ò
```

```
;
                                                                                                                                                                                                                                                                                                                                                                                        614
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          112 CGSSDLYLVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAV 171
                                                                                                                                                                                                                                                                                                        51
                                                                                                                                                                                                                                                                                                   3 KKGSVVIVG---RIVLNG-----AYAQQIRGEEGCQETSQIGRDKNQVEGEVQIVST
                                                                                                                                                                                                                                                                                                                                  MEDLINE-21262212; PubMed-11369872;
Lanford R.E., Lee H., Chavez D., Guerra B., Brasky K.M.;
"Infectious cDNA clone of the hepatitis C virus genotype 1 prototype
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        K
                                                                                                                                                                                                                                                             Gaps
PROSITE; PS50521; RDRP_VIRAL; 1.
ATP-binding; Coat protein; Envelope protein; Glycoprotein; Helicase; Hydrolase; Nonstructural protein; Polyprotein; RNA-directed RNA polymerase; Transmembrane.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 sequence..;
J. Gen. Virol. 82:1291-1297(2001).
J. CAUDAROTEIN BUVELOPE CONSISTS OF TWO PROTEINS:
PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MRNA (BY SIMILARITY).
ENBL. AF211632; AAF81759.1;
J. HSSP: P27958; JALV.
InterPro; JPR000345; CYtC_heme_bind.
InterPro; JPR0001410; DEAD.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
Hepacivirus.
                                                                                                                                                                                                              DB 12; Length 2436;
                                                                                                                                                                                                                                                           Indels
                                                                                                                                                                 2436 AA; 264734 MW; D7B9872900BE3125 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     09IFE5;
0-1-CTT-2000 (TrEMBLrel. 15, Created)
01-CCT-2000 (TrEMBLrel. 15, Last sequence update)
01-WAR-2003 (TrEMBLrel. 23, Last annotation update)
                                                                                                                                                                                                                                 l.3e-82;
hes 10;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    PRI; 3011 AA.
                                                                                                                                                                                                                                        Pred. No. 1.3e
8; Mismatches
                                                                                                                                                                                                              87.6%; Score 894.5; 85.8%; Pred. No. 1.3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          172 CIRGVAKAVDFIPVESLETIMRSP 195
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  InterPro: IPR002868; HCV_NSS.
InterPro: IPR001166; HCV_RARP.
InterPro: IPR001169; HcV_RARP.
InterPro: IPR001095; RNA_pol_bS_PS.
InterPro: IPR007094; RNA_pol_PSvir.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              IPR002522; HCV_capsid.
IPR002521; HCV_core.
IPR002519; HCV_core.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                HCV_NS1.
HCV_NS2.
HCV_NS3.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Pfam; PF01543; HCV_capsid; 1. pfam; PF01542; HCV_core; 1. Pfam; PF01539; HCV_env; 1. Pfam; PF01560; HCV_NS1; 1.
                                                                                                                                                                                                                                   Best Local Similarity 85.8
Matches 175; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      PRELIMINARY;
                                                                                                                                              2436
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              IPR002531; H
IPR002518; H
IPR004109; H
IPR000745; H
IPR001490; H
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Genome polyprotein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      NCBI_TaxID=11103;
                                                                                                                           1
2436
```

g ò QQ ò a ò g

```
1005 RKGREILLGPADGMVSRGWRLLAPITAYAQQTRGLLGCIITSLTGRDKNQVEGEVQIVST 1064
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               52 AAQTFLATCINGVCWTVYHGAGTRTIASPKGPVIQMYTNVDKDLVGWPAPQGSRSLTPCT 111
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  112 CGSSDLYLVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAV 171
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    3 KKGSVVIVG---RIVLNG----AYAQOTRGEEGCQETSQTGRDKNQVEGEVQIVST
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        11; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               PROSITE; PS50507; RDRP_POSITIVE; 1.
PROSITE; PS50521; RDRP_VIRAL; 1.
AIP-binding; Coat protein; Envelope protein; Giycoprotein; Helicase;
              *Full-length nucleotide sequence of a Japanese hepatitis C virus isolate (HC-Jl) with high homology to USA isolates."; Nucleic Acids Res, 20:6410-6410(1992).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    DB 12; Length 3011;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               "Characterization and mapping of a B-cell immunogenic domain in
                                                                                                                                                                                                                                                                                                                                                                                        MEDLINE-94174722; PubMed-7510436;
Mink M., Benichou S., Madaule P., Tiollais P., Prince A.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 11; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Hydrolase: Nonstructural protein; Polyprotein;
RNA-directed RNA polymerase; Transferase; Transmembrane.
SEQUENCE 3011 AA; 327112 MW; 97E9052C0250463B CRC64;
                                                                                                                                                                                                                                                        Submitted (DEC-1992) to the EMBL/GenBank/DDBJ databases
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             87.5%; Score 893.5; DB 1:
85.8%; Pred. No. 2.2e-82;
tive 7; Mismatches 11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          InterPro: IPR001650; Helicasc.C.
InterPro: IPR001650; RNA_pol_DS_PS.
InterPro: IPR001095; RNA_pol_DS_PS.
InterPro: IPR001094; RNA_pol_DS_PS.
Pfam: PF01540; HCV_capsid; 1.
Pfam: PF01542; HCV_core; 1.
Pfam: PF01550; HCV_core; 1.
Pfam: PF01530; HCV_NS1; 1.
Pfam: PF02907; HCV_NS3; 1.
Pfam: PF01506; HCV_NS3; 1.
Pfam: PF01001; HCV_NS4a: 1.
Pfam: PF01506; HCV_NS4a: 1.
Pfam: PF01506; HCV_NS5a: 1.
Pfam: PF01509; HCV_NS4a: 1.
Pfam: PF01509; HCV_NS4a: 1.
Pfam: PF01509; HCV_NS5a: 1.
Pfam: PF01509; HCV_NS5a: 1.
Pfam: PF01509; Viral_RAPP: 1.
Pfam: PF0180627; HCV_NS5a: 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 HSSP; PZ/930; inti.
HSSP; PZ/930; interPro; IPR001410; DEAD.
InterPro; IPR00252; HCV_capsid.
InterPro; IPR002521; HCV_core.
InterPro; IPR002531; HCV_env.
InterPro; IPR002519; HCV_MS1.
InterPro; IPR004109; HCV_MS3.
InterPro; IPR00445; HCV_NS3.
InterPro; IPR001490; HCV_NS4.
InterPro; IPR001490; HCV_NS4.
InterPro; IPR001490; HCV_NS4.
InterPro; IPR001460; HCV_NS5a.
InterPro; IPR001668; HCV_NS5a.
InterPro; IPR001668; HCV_NS5a.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Ouery Match
Best Local Similarity 85.84
Matches 175; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             InterPro; IPR002868;
InterPro; IPR002166;
InterPro; IPR001650;
                                                                                                                                                                                                                                                                                                                        SEQUENCE FROM N.A.
                                                                                                                                                     SEQUENCE FROM N.A.
                                                                                                                                                                                        STRAIN-HC-J1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                               Inchauspe G.;
                                                                                                                                                                                                                        Okamoto
                 SOW WE WANT THE PROPERTY OF TH
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  å
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CGSSDLYLVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAV 171
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      21
R Pfam; PF01538; HCV_NS2; 1.

R Pfam; PF01006; HCV_NS3; 1.

R Pfam; PF01006; HCV_NS4s; 1.

R Pfam; PF01006; HCV_NS4s; 1.

R Pfam; PF01001; HCV_NS4s; 1.

R Pfam; PF00998; Viral_RRP; 1.

R ProDom; PD186062; HCV_NS1; 1.

R ProDom; PD186062; HCV_NS1; 1.

R RNGATE; PS05090; CTTCCHROME_C; 1.

R PROSITE; PS0501; RDRP_POSTTIVE; 1.

R PROSITE; PS50521; RDRP_VIRAL; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Okamoto H., Okada S., Sugiyama Y., Kurai K., Iizuka H., Machida A., Miyakawa Y., Mayumi M.;
"Nucleotide sequences of the genomic RNA of hepatitis C virus isolated from a human carrier: comparison with reported isolates for conserved and divergent regions.";
J. Gen. Virol. 72:2697-2704(1991).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      3 KKGSVVIVG---RIVLNG-----AYAQOTRGEEGCQFTSOTGRDKNQVEGEVQIVST
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
Hepacivirus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      11;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   DB 12; Length 3011;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          "The 5'-terminal sequence of the hepatitis C virus genome.";
Jpp. J. Exp. Med. 60:167-177(1990).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Okamoto H., Okada S., Sugiyama Y., Yotsumoto S., Tanaka T.,
Yoshizawa H.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               87.6%; Score 894.5; DB 12; Length
85.8%; Pred. No. 1.7e-82;
iive 8; Mismatches 10; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   2489CE74AC864E58 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                            Hydrolase; Nonstructural protein; Polyprotein; RNA-directed RNA polymerase; Transferase; Transmembrane SEQUENCE 3011 AA; 327124 MW; 2489CE74AC864E58 CRC64.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Created)
Last sequence update)
Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            PRT; 3011 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        172 CTRGVAKAVDFIPVESLETTMRSP 195
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               STRAIN-HC-J1;
MEDLINE-93117120; PubMed-1335573;
Okamoto H., Kanal N., Mishiro S.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               STRAIN-HC-J1;
MEDLINE-91013116: PubMed-2170712;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 MEDLINE-92044440; PubMed=1658196;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        01-NOV-1996 (TrEMBLrel. 01, 01-NOV-1996 (TrEMBLrel. 01, 01-MAR-2003 (TrEMBLrel. 23, Genome polyprotein. Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Local Similarity 85.8 les 175; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               NCBI_TaxID-11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 STRAIN-HC-J1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 112
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        003463
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            RESULT 4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            DOT ACCOUNT OF THE READ OF THE
```

```
Pfam: PF01001; HCV_NS4b; 1.
Pfam: PF01506; HCV_NS5a; 1.
Pfam: PF00271; heliase_C; 1.
Pfam: PF00998; Viral_RGFP; 1.
ProDom; PD186062; HCV_NS1; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        EMBL; AF009606; AAB66324.1;
HSSP; P27958; 1HEI.
InterPro; IPR001410; DEAD.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                transcribed RNA.";
Science 277:570-574(1997).
                                                                Conservative
                                                                                                                                                                                                                                                                               PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                           Genome polyprotein.
                                                    Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                     SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                          Hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                               NCBI_TaxID=11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          PF01006;
                                                               Matches 171;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      InterPro;
InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               InterPro;
InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    nterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          interPro;
               SEQUENCE
                                       Query Match
Best Local
                                                                                                                                                                                                                                                                                        036579;
                                                                                                                                                                                                                                                                               036579
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Pfam;
                                                                                                                                                                                                                                                       RESULT 7
                                                                                                                                                                                                                                                                                            F.1
SO
                                                                                                              q
                                                                                                                                                                g
                                                                                                                                                                                          ò
                                                                                                                                          ö
                                                                                                                                                                                                                                                                                                                                                                                                                                  ပွဲ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           78 ASPKGPVIQMYINVDKDLVGWPAPQGSRSLIPCTCGSSDLYLVIRHADVIPVRRRGDSRG 137
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        65 ASPKGPVIQMYINVDKDLVGWPAPQGSRSLTPCTCGSSDLYLVTRHADVIPVRRRGDSRG 124
                                                                                                                                                                                                                                                                                                                                                                                                                                                           17
                                                                                                                                                                                                                                                                                                                                                                                                                                                                    138 SLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVCTRGVAKAVDFIPVESLETTMRS 194
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        18 AYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVSTAAQTFLATCINGVCWTVYHGAGTRTI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Holland-Staley C.A., Kovari L.C., Golenberg E., Mayers D.L.;
"Genetic Diversity and response to IFN of the NS3 Protease Gene from Clinical Strains of the Hepatitis C Virus.";
Submitted (ARR-2001) to the FMBL/GenBank/DDBJ databases.
EMBL, AF36218; AAK55541.;
InterPro: IPR004109; HCV_NS3.
                                                                                                                                                                                                                         SEQUENCE FROM N.A.
STRAIN-PL.IY:

Holland-Staley C.A., Kovari L.C., Golenberg E., Mayers D.L.;

"Genetic Diversity and response to IFN of the NS3 Protease Gene from Clinical Strains of the Hepatitis C Virus.";

EMBL, SENDITED (APR-2001) to the EMBL/GenBank/DDBJ databases.

EMBL, AR362335; AAR54560.1;

InterPro: IPR004109; HCV.NS3.

Protease.
                                                                                                                                                                                                                                                                                                                                                                                                                                 0; Gaps
                                                                                                                                                                   viruses; ssRNA positive-strand viruses, no DNA stage: Flaviviridae;
Hepacivirus.
NCBI_TaxID=11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 VI uses: ssrnA positive-strand viruses, no DNA stage; Flaviviridae;
Hepacivirus.
                                                                                                                                                                                                                                                                                                                                                                                                   Score 891; DB 12; Length 181;
Pred. No. 1.1e-83;
1; Mismatches 5; Indels (
                                                                                                                                                                                                                                                                                                                                                                    181 181
181 AA; 19130 MW; 85D91869299B7C35 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         01-DEC-2001 (TrEMBLrel. 19, Created)
01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
                                                                                                              01-DEC-2001 (TrEMBLrel. 19, Created)
01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  181 AA
                                                                                        181 AA
             172 CTRGVAKAVDFIPVESLETIMRSP 195
                                                                                        PRT;
                                                                                                                                                                                                                                                                                                                                                                                                      Query Match
Best Local Similarity 96.6%;
Matches 171; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               NS3 protease (Fragment).
Hepatitis C virus.
                                                                                                                                                      NS3 protease (Fragment). Hepatitis C virus.
                                                                                        PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 NCBI_TaxID=11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        STRAIN-Pt . 4;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Protease.
NON_TER
                                                                                                                                                                                                                                                                                                                                                       NON_TER
NON_TER
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 091RT5
                                                                                        Q91RR8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         RESULT 6
                                                               RESULT 5
Q91RR8
```

g

Q ò

ò

```
0;
                                                                                                                                                                                                                                                                                78 ASPKGPVIQMYTNVDKDLVGWPAPQGSRSLIPCTCGSSDLYLVTRHADVIPVRRRGDSRG 137
                                                                                                                                                                                                                                                                                                          18 AYAQQTRGEEGCQETSQTGRDKNQVEGEYQIVSTAAQTFLATCINGVCWTVYHGAGTRTI 77
                                                                                                                                                                                                  5 AYAQQTRGLLGCIITSLIGRDKNQVEGEVQIVSTAAQTFLATCINGVCWTVYHGAGTRII 64
                                                                                                                                                                                                                                                                                                                                                                                     Kolykhalov A.A., Agapov E.V., Blight K.J., Mihalik K., Feinstone S.M.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          -!- SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS: PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MRNA (BY SIMILARITY).
                                                                                                              Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae; Hepacivirus.
                                                                                                              ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               "Transmission of hepatitis C by intrahepatic inoculation with
                                                       Length 181;
                                                                                                              Indels
181 AA; 19130 MW; 85D91869299B7C35 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         01-JAN-1998 (TrEMBLrel. 05, Created)
01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
                                                    87.3%; Score 891; DB 12; 96.6%; Pred. No. 1.1e-83;
                                                                                                           1; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     PRT; 3011 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          InterPro; IPR007095; RNA_pol_DS_PS.
InterPro; IPR007094; RNA_pol_PSvir.
Pfan: PP01543; HCV_capsid; 1.
Pfan: PF01542; HCV_core: 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            | IPR002521 | HCV_capsid. |
| IPR002521 | HCV_capsid. |
| IPR002521 | HCV_core. |
| IPR002531 | HCV_NS1. |
| IPR002518 | HCV_NS2. |
| IPR004109 | HCV_NS3. |
| IPR001490 | HCV_NS3. |
| IPR001490 | HCV_NS4. |
| IPR00166 | HCV_NS4. |
| IPR00166 | HCV_NS4. |
| IPR00166 | HCV_NS4. |
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   MEDLINE-97373636; PubMed-9228008
```

S

```
96.08;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Matches 170; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              NS3 protease (Fragment).
Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             NCB1_TaxID-11103;
                                                                                                                                                                                                                                                                                    SEQUENCE FROM N.A.
                                                                                                                                                                                                                                       NCBI_TaxID-11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    181
                                                                                                                                                                                                                                                                                                           STRAIN-Pt.4B;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    STRAIN-Pt.K;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Hepacivirus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  NON_TER
NON_TER
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   NON_TER
NON_TER
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Protease.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Best Local
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Q91RS1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          RESULT 10
Q91RS1
RESULT 9
                      091RR3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             õ
                                                                                                                                                                                                                                                                                                                                                                          2
                                                                                                                                                                                                                                                                                      78 ASPKGPVIQMYINVDKDLVGWPAPQGSRSLIPCTCGSSDLYLVTRHADVIPVRRRGDSRG 137
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           65 ASPKGPVIQMYTNVDQDLVGWPAPQGARSLTPCTCGSSDLYLVTRHADVIPVRRKGDSRG 124
                                                                                                                                                                                                                                                                                                                                                 52 AAQTFLATCINGVCWTVYHGAGTRIIASPKGPVIOMYINVDKDLVGWPAPQGSRSLTPCT 111
                                                                                                                                                                                                                                                                                                                                                                                                                                        112 CGSSDLYLVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAV 171
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     5 AYAQOTRGLLGCIXISLIGRDKNOVEGEVQIVSIAAQIFLAICINGVCWIVIHGAGIRII 64
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVCTRGVAKAVDFIPVESLETIMRS 194
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               18 AYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVSTAAQTFLATCINGVCWTVYHGAGTRTI
                                                                                                                                                                                                                                                                3 KKGSVVIVG---RIVLNG----AYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVST
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Gene from
                                                                    Helicase;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    .;
0
                                                                                                                                                                          DB 12; Length 3011;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Length 181;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Holland-Staley C.A., Kovari L.C., Golenberg E., Mayers D.L. Genetic Diversity and response to IFN of the NS3 protease (Clinical Strains of the Hepatitis C Virus."; Submitted (APR-2001) to the EMBL/GenBank/DDBJ databases. EMBL; AFS59238; AAK78563.1; "MEML; AFS69238; AAK78553.1; "PRO1907; HCV_NS3: 1.
                                                               ATP-binding; Coat protein; Envelope protein; Glycoprotein; Hydrolase; Nonstructural protein; Polyprotein; RNA directed RNA polymerase; Transferase; Transsmembranc. SEQUENCE 3011 AA; 327182 MW; E2E0EE809c63c1B9 CKC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Indels
                                                                                                                                                                                                                     11; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           181 181
181 AA; 19084 MW; 3B5E816IF2100A72 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        01-DEC-2001 (TrEMBLrel. 19, Created)
01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               'n
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      87.1%; Score 889; DB 12;
96.0%; Pred. No. 1.8e-83;
Live 2; Mismatches 5;
                                                                                                                                                                          87.1%; Score 889.5; DB 1:
84.8%; Pred. No. 5.5e-82;
ive 9; Mismatches 11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1185 CTRGVAKAVDFIPVENLETTMRSP 1208
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CTRGVAKAVDFIPVESLETTMRSP 195
                    PROSITE; PS50507; RDRP_POSITIVE; 1. PROSITE; PS50521; RDRP_VIRAL; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Conservative
                                                                                                                                                                                                                     Matches 173; Conservative
SM00487; DEXDC; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         NS3 protease (Fragment).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     01-DEC-2001 (TrEMBLrel.
01-DEC-2001 (TrEMBLrel.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Similarity
                                                                                                                                                                                                   Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           NCBI_TaxID=11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               STRAIN-Pt.3U;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Local Simi
nes 170;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Hepacivirus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Protease.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           172
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             138
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Query Match
                                                                                                                                                                              Query Match
                                                                                                                                                                                                     Sest Local
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 091RR5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               RESULT 8
091RR5
1001-D
DT 001-D
DT 001-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                ö
                                                                                                                                                                                                                                                                                                      a
                                                                                                                                                                                                                                                                                                                                                 ç
                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                              셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           οχ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  a
```

```
78 ASPRGPVIQMYINVDKDLVGWPAPQGSRSLTPCTCGSSDLYLVTRHADVIPVRRRGDSRG 137
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           65 ASPKGPVIQMYINVDRDLVGWPAPQGARSLIPCTGGSSDLYLVTRHADVIPVRRRGDSRG 124
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           64
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         5 AYAQQTRGLLGCIIISLTGRDKNQYEGEVQIVSTAAQTFLATCINGYCHTVYHGAGTRII
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     18 AYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVSTAAQTFLATCINGVCWTVYHGAGTRTI
                                                                                                                                                                                                                                                                                                                                                                    Holland-Staley C.A., Kovari L.C., Golenberg E., Mayers D.L.;
"Genetic Diversity and response to IfN of the NS3 Protease Gene from
"Genetic Diversity and response to IfN of the NS3 Protease Gene from
Clinical Strains of the Hepatitis C Virus.";
Submitted (APR-2001) to the EMBL/GenBank/DDBJ databases.

EMBL, AF865401, AR85465.1;
-.
Interpro, IPR004109; HCV_NS3.
                                                                                                                                                                         Heparitis C virus.
Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
Hepacivirus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Gene
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ô
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Length 181;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Holland-Staley C.A., Rovari L.C., Golenberg E., Mayers D.L. Genetic Diversity and response to TFN of the NS3 Protease of Clinical Strains of the Hepatitis C Virus."; Submitted (APR-2001) to the EMBL/GenBank/DDBJ databases. EMBL, AF856232, AAK54557.1; "Interpro: IPRO4109; HCV-NS3.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          5; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               181 AA; 19114 MW; ABB90B5B3ABA4E26 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 181 181
181 AA; 19115 MW; 5D85F88AD7AC1All CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  01-DEC-2001 (TrEMBLrel. 19, Created)
01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
                                                                                         Last sequence update)
Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Score 888; DB 12;
Pred. No. 2.3e-83;
2; Mismatches 5;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Score 888; DB 12;
Pred. No. 2.3e-83;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    87.0%; Score 888;
                                                                 Created)
Q91RR3 PRELIMINARY;
Q91RR3;
Q1-DEC-2001 (TrEMBLrel. 19, Cr
Q1-DEC-2001 (TrEMBLrel. 19, La
Q1-MAR-2003 (TrEMBLrel. 23, La
NS3 protease (Fragment).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     87.0%;
96.0%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Ouery Match
Best Local Similarity
```

ö

```
SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           NCBI_TaxID=11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            STRAIN-Pt.3T;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  NON_TER
NON_TER
SEQUENCE
                                                                                                                                                                               Protease.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Protease.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Query Match
                                                                                                                                                                                                                                           SEQUENCE
                                                                                                                                                                                                                                                                                        Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Q91RR6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    RESULT 13
                                                                                                                                                                                                                                                                                      Query
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          091RR6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    a
    REPRESENT NO DE NO DE NET TAN DE 
                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ô
    ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          65 ASPKGPVIQMYINVDKDLVGWPAPQGARSLIPCTCGSSDLYLVTRHADVIPVRRKGDSRG 124
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   78 ASPKGPVIQMYINVDKDLVGWPAPQGSRSLTPCTCGSSDLYLVTRHADVIPVRRRGDSRG 137
                                                                                                                                  ASPKGPVIQMYTNVDKDL/VGWPAPQGSRSLTPCTCGSSDL/YL/VTRHADVIPVRRRGJSRG 137
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         5 AYAQOTRGLIGCIITSLIGRDKNOVEGEVQIVSTAAQIFLATCINGVCWTVYHGAGIRTI 64
                                               11
                                                                 5 AYAQOTRGLLGCIITSLTGRDKNQVEGEVQIVSTAAQTFLAICINGVCWTVYHGAGTRTI 64
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      138 SLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVCTRGVAKAVDFIPVESLETTMRS 194
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               138 SLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVCTRGVAKAVDFIPVESLETTMRS 194
                                                                                                                                                                                                                                           18 AYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVSTAAQTFLATCINGVCWTVYHGAGTRTI
                                               AYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVSTAAQTFLATCINGVCWTVYHGAGTRTI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Folland-Staley C.A., Kovari L.C., Golenberg E., Mayers D.L.;
"Genetic Diversity and response to IFN of the NS3 Protease Gene from Clinical Strains of the Hepatitis C Virus.";
Submitted (APR-2001) to the EMBL/GenBank/DDBJ databases.
EMBL, AF369245; AAK4570.1:
InterPro; IPR004109; HCV_NS3.
Pfam; PF02907; HCV_NS3; 1.
      Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
    ..
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           87.0%; Score 888; DB 12; Length 181; 96.0%; Pred. No. 2.3e-83; tive 2; Mismatches 5; Indels (
    Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Q91RQ8;
01-DEC-2001 (TrEMBLrcl. 15, Created)
01-DEC-2001 (TrEMBLrcl. 15, Last sequence update)
01-MAR-2003 (TrEMBLrcl. 23, Last annotation update)
NS3 protease (Fragment).
Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                01-DEC-2001 (TrEMBLrel. 19, Created)
01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
    5.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             181 AA
                                                                                                                                                                                                                                                                                                                                                                         181 AA
  2; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Conservative
    Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               NS3 protease (Fragment).
Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                    PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                [1]
SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               NCBL_TaxID-11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         NCBI_TaxID-11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                STRAIN-Pt.52;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Hepacivirus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Best Local Sim
Matches 170;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Hepacivirus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Protease.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          NON_TER
NON_TER
SEQUENCE
                                                                                                                               28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Q91RT1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          091RT1
                                                                                                                                                                                                                                                                                                                                                                       091RQ8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             RESULT 12
091RT1
10 091RT1
00 091RT
DT 01-DE
DT 01-DE
DT 01-MS3 p
OS Hepat
OC Hepac
OX NCBI_
RP, SEQUE
                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                               g
                                                                                셤
                                                                                                                               ò
                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                             셤
                                                 ò
```

```
ö
                                                                                                                                                                                                                                                                                                                                   ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                78 ASPKGPVIQMYTNVDKDLVGWPAPQGSRSLTPCTCGSSDLYLVTRHADVIPVRRRGDSRG 137
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        78 ASPKGPVIQMYTNVDKDLVGWPAPQGSRSLTPCTCGSSDLYLVTRHADVIPVRRRGDSRG 137
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 77
                                                                                                                                                                                                                                                                                                                                                                                                   5 AYAQQTRGLLGCIITSLTGRDKNQVEGEVQIVSTAAQTFLATCINGVCWTVYHGAGTRII 64
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               18 AYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVSTAAQTFLATCINGVCWTVYHGAGTRTI 77
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            64
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     5 AYAQQIRGLLGCIITSLTGRDKNQVEGEIQIVSTAAQIFLATCINGVCWIVYHGAGIRII
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   138 SLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVCTRGVAKAVDFIPVESLETIMRS 194
                                                                                                                                                                                                                                                                                                                                                                              18 AYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVSTAAQTFLATCINGVCWTVYHGAGTRTI
"Genetic Diversity and response to IFN of the NS3 Protease Gene from Clinical Strains of the Hepatitis C Virus."; Submitted (APR-2001) to the EMBL/GenBank/DDBJ databases. EMBL, AF369222; AAK51547.1; InterPro; IPR004109; HCV_NS3. Pfam. PF02907; HCV_NS3; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Gene from
                                                                                                                                                                                                                                                                                                                                      Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
Hepacivirus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ö
                                                                                                                                                                                                                                                                                                                                 ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Holland-Staley C.A., Kovari L.C., Golenberg E., Mayers D.L.;
"Genetic Diversity and response to IFN of the NS3 Protease Ge
Clinical Strains of the Hepatitis C Virus.";
Submitted (APR-2001) to the EMBL/GenBank/DDBJ databases.
EMBL, ARS62337, AAKS6452.1;
InterPro; IPR004109; HCV_NS3.
Pfam; PF02907; HCV_NS3; I.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Length 181;
                                                                                                                                                                                                                                                                                  Length 181;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Score 886; DB 12; Length 10
Pred. No. 3.7e-83;
3; Mismatches 5; Indels
                                                                                                                                                                                                                                                                                                                                 Indels
                                                                                                                                                                                                               181 181
181 AA: 19114 MW; ABB90B5B3ABA4E26 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 181 AA; 19101 MW; 614ADA8B0F33CCAF CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            01-DEC-2001 (TrEMBLrel. 19, Created)
01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
01-NAR-2003 (TrEMBLrel. 23, Last annotation update)
                                                                                                                                                                                                                                                                                    Score 888; DB 12;
Pred. No. 2.3e-83;
                                                                                                                                                                                                                                                                 87.0%; Scur.
96.0%; Pred. No. 2...
2; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  3; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     86.8%;
95.5%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Best Local Similarity 95.5
Matches 169; Conservative
                                                                                                                                                                                                                                                                                                         Best Local Similarity 96.0
Matches 170; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                NS3 protease (Fragment).
```

```
Job time : 38.2105 secs
                         InterPro;
a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                              78 ASPKGPVIQMYINVDKDLVGWPAPQGSRSLIPCICGSSDLYLVIRHADVIPVRRRGDSRG 137
                                                                                                                                                                                                                                                                                                                                                                                                                                                           The STRAIN-Colone;

STRAIN-Colone;

Desai S.M., Devare S., Yamaguchi J.;

The partitis C Virus.";

"Hepartitis C Virus.";

"Hepartitis C Virus.";

"Hepartitis C Virus.";

Submitted (Jul.-2000) to the EMBL/GenBank/DDBJ databases.

Lishburg: THE VIRUON OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A LIPOPROTEIN ENVELOPE CONSISTS OF TWO PROTEINS:

PROTEIN AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN APPONDRY;

RESEN: AF290978; AAG02099-1; ...

RESSP; P27958; HREI.

INTERPRO: IPRO01410; DEAD.

RICHEPRO: IPRO015221; HCV_cose.
                                                                                                                                                                                                                                                                                                                                                                                               77
                                                                                                                                                                                                                                                                                                                                                                                                           138 SLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVCTRGVAKAVDFTPVESLETTMRS 194
                                                                                                                                                                                                                                                                                                                                                                                             18 AYAQOTRGEEGCQETSQTGRDKNQVEGEVQIVSTAAQIFILATCINGVCWTVYHGAGTRTI
                                                                                                                                                                                 STRAIN-PL.174;
Holland-Staley C.A., Kovari L.C., Golenberg E., Mayers D.L.;
Holland-Staley C.A., Kovari L.C., Golenberg E., Mayers D.L.;
Genetic Diversity and response to IFN of the NS3 Protease Gene from Clinical Strains of the Hepatitis C Virus.";
Submitted (APR-2001) to the EMBL/GenBank/DDBJ databases.
EMBL; AF369224; AAK54549.1;
InterPro; IPR004109; HCV_NS3.
                                                                                                                                                                                                                                                                                                                                                                         Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                        Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                         ö
                                                                                                                                                                                                                                                                                                                                                Length 181;
                                                                                                                                                                                                                                                                                                                                                                         5; Indels
                                                                                                                                                                                                                                                                                                            181 181
181 AA; 19131 MW; 8BD7FC2769DBD635 CRC64;
                                                           01-DEC-2001 (TrEMBLrel. 19, Created)
01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Last sequence update)
Last annotation update)
                                                                                                                                                                                                                                                                                                                                               86.8%; Score 886; DB 12; 96.0%; Pred. No. 3.7e-83;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  PRT; 3011 AA
                                     181 AA
                                                                                                                                                                                                                                                                                                                                                                       2; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Created)
                                     PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          01-MAR-2001 (TrEMBLrel. 16, 01-MAR-2001 (TrEMBLrel. 16, 01-MAR-2003 (TrEMBLrel. 23,
                                                                                                                                                                                                                                                                                                                                                             Best Local Similarity 96.0
Matches 170; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    PRELIMINARY;
                                                                                                  protease (Fragment).
titis C virus.
                                     PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Genome polyprotein.
Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SEQUENCE FROM N.A.
                                                                                                                                                                [1]
SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               NCBI_TaxID=11103;
                                                                                                                                                NCBI_TaxID-11103;
                                                                                                               Hepatitis C
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Hepacivirus
                                                                                                                                   Hepacivirus
                                                                                                                                                                                                                                                                                    Protease.
NON_TER
NON_TER
                                                                                                                                                                                                                                                                                                                          SEQUENCE
                                                                                                                                                                                                                                                                                                                                                 Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Q9ELS8
                                    Q91RS9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           RESULT 15
Q9ELS8
           RESULT 14
091RS9
                                                                                                   NS3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    QQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                      a
                                                                                                                                                                                                                                                                                                                                                                                                                                              à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
```

```
52 AAQTFI.ATCINGVCWTVYHGAGTRIIASPKGPVIQMYTNVDKDLVGWPAPQGSRSLTPCT 111
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        3 KRGSVVIVG --- RIVLNG --- -- AYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVST
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   11; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         PROSITE; PS00190; CTYCCHRONE_C; 1.
PROSITE; PS00190; CTYCCHRONE_C; 1.
PROSITE; PS50507; RORP_POSITIVE; 1.
PROSITE; PS50507; RORP_VIRAL; 1.
PROSITE; PS50501; RORP_VIRAL; 1.
Hydrolase; Psofton: Broelope protein; Glycoprotein; Helicase; Hydrolase; Nonstructural protein; Polyprotein;
RNA-directed RNA polymerase; Transferase; Transmembrane.
SEQUENCE 3011 AA: 327107 MW; AGBECFSABBEE13F CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Length 3011;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                11; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        DB 12;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Query Match

86.7%; Score 885.5; DB 17

Best Local Similarity 84.3%; Pred. No. 1.4e-81;

Matches 172; Conservative 10; Mismatches 11.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1185 CTRGVAKAVDFIPVENLETTMRSP 1208
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Search completed: August 30, 2003, 19:18:18
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 172 CTRGVAKAVDFIPVESLETTMRSP 195
                                                 Interpro: IPR004109; HCV_NS3.
Interpro: IPR000445; HCV_NS4a.
Interpro: IPR001040; HCV_NS4a.
Interpro: IPR0012868; HCV_NS5a.
Interpro: IPR001266; HCV_RSARP.
Interpro: IPR0010560; Helicase_C.
Interpro: IPR0010560; Helicase_C.
Interpro: IPR007095; RNA_pol_DS_PS.
Interpro: IPR0071095; RNA_pol_DS_PS.
Pfam: PF01543; HCV_Copeid: 1.
                                                                                                                                                                                                                                                                                                                             Pfam, PF01539; HCV_ENV; 1. Pfam; PF01539; HCV_ENV; 1. Pfam; PF01539; HCV_ENS; 1. Pfam; PF02907; HCV_ENS; 1. Pfam; PF01001; HCV_ENS4; 1. Pfam; PF01001; HCV_ENS4; 1. Pfam; PF01505; HCV_ENS4; 1. Pfam; PF01505; HCV_ENS5; 1. Pfam; PF00571; helicase_C; 1. Pfam; PF00998; VITAL_RGRP; 1. Pr0Dom; PD186062; HCV_ENS5; 1. SWART; SW00487; DEXDC; 1.
HCV_NS1.
IPR002531;
IPR002518;
```

ñ

```
ARI18692 Sequence 10440 Sequence 54 I009329 Sequence 109329 Sequence 109331 Sequence 108331 Sequence 108331 Sequence 108331 Sequence 108294 Sequence 171894 Sequence 171894 Sequence 9 FR17689 Sequence 9 M63321 Hepatitis C AF397808 Synthetic AF397805 Synthetic AF397807 Synthetic AF397807 Synthetic AF397807 Synthetic
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           D10749 Hepatitis C
AR127809 Sequence
BD081910 Hepatitis
AR127810 Sequence
BD081911 Hepatitis
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ARIA5264 Sequence
AF511948 Hepatitis
AF369218 Hepatitis
AF369235 Hepatitis
AX441176 Sequence
AX467113 Sequence
                                                                                                                                                                                                                                                                                   AF268278 Pestiviru
AR118686 Sequence
IO6434 Sequence 48
IO9328 Sequence 8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            AX395309 Sequence
AX454018 Sequence
E08263 GRNA of Hep
E08264 CDNA of Hep
D10749 Hepatitis C
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     gRNA of Hep
CDNA of Hep
                                                                                                                                                                                                                                                                     AR179057 Sequence
                                                                                                                                                                Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.
                                                                                                                                                                                                                                                 Description
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ALIGNMENTS
                                                                                                                                                                                                                SUMMARIES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        HPCPLYPRE
AF387805
AF387808
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       AF387806
AF387807
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    AR301300
AR176483
BD080334
E66593
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               AF369218
                                                                                                                                                                                                                                                                                                                                                                                                                  AR118728
AR118722
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          AF369235
                                                                                                                                                                                                                                                                                               AR118686
                                                                                                                                                                                                                                                                                                                                                                   AR118696
                                                                                                                                                                                                                                                                                                                                                                                                                                                    BD091382
                                                                                                                                                                                                                                                                                                                                                                                                                                                                         AR166930
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   I71894
I81885
                                                                                                                 em_htgo_hum:*
em_htgo_mus:*
em_htgo_other:*
                      .htg_inv:*
.htg_other:*
                                                       em_htg_pln: *
em_htg_rod: *
em_htg_mam: *
           em_htg_hum:*
                                              em_htg_mus:*
                                                                                           em_htg_vrt:*
                                                                                                                                                                                                                          %
Query
Match Length DB
                                                                                                                                                                                                                                                                                                                                6785
6785
6785
7310
7310
7310
8310
8316
8987
9185
9185
9185
9185
9401
9401
9401
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              9401
9401
9609
9609
9618
9646
                                                                                                       em_sy:
                        5
                                  Ē
                                                                                                                                                                                                                                                                                                                                           Score
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            RESULT 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                    Command line parameters:
-MODEL-frame+pp. modcl -DEV-xlp
-Q-/cgp12_1/0SPTO_spool/US09965594/runat_29082003_151919_28310/app_query.fasta_1.2872
-Q-/cgp12_1/OSPTO_spool/US09965594/runat_29082003_151919_28310/app_query.fasta_1.2872
-Q-/cgp12_1/OSPTO_spool/US09965594/runat_29082003_1.CODPCL-0.LOOPCL-0.LOOPEXT-0
-UNITS-bits.START-1.END--1.MATRIX-biosum62 -TRANS-human40.cdi -LISI-45
-UOTALIGN-200 -THR_SCORE-CT -THR_MAX-100 -THR_MIN-0.ALGN-15.MODE-LOCAL
-USFR-USO9965594_QCGN_1 | 1.4686_Q*runat_29082003_151919_28310 -NCPU-6 -TCPU-3
-NO_WMAP -LARGEQUERY NEC_SCORES-0.WAIT -DSPBLOCK-100 -LONGLOC
-DBV_INBOUT-120 -WARN TIMEOUT-30 -THERADS-1 -XORPOP-10 -XGAPEXT-0.5 -FGAPOP-6
                                                                                           (without alignments)
3147.423 Million cell updates/sec
                                                                                                                                                    1 MKKKGSVVIVGRIVLNGAYA.......VAKAVDFIPVESLETIMRSP 195
                                                                              August 30, 2003, 19:18:33 ; Search time 2534.57 Seconds
       GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.
                                                        - nucleic search, using frame_plus_p2m model
                                                                                                                                                                                                                                                                     Total number of hits satisfying chosen parameters:
                                                                                                                                                                                                                                               2888711 scqs, 20454813386 residues
                                                                                                                                                                                                                                                                                                                              Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries
                                                                                                                                                                      BLOSUM62
Xgapop 10.0 , Xgapext /
Ygapop 10.0 , Ygapext /
Fgapop 6.0 , Fgapext Delop 6.0 , Delext
                                                                                                                                                                                                                                                                                            Minimum DB seq length: 0
Maximum DB seq length: 200000000
                                                                                                                           US-09-965-594-12
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            9D_0v:*
9D_ph:*
9D_ph:*
9D_pr:*
9D_pr:*
9D_pr:*
9D_sy:*
9D_sy:*
9D_un:*
9D_vi:*
em_fun:*
em_fun:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               em_pat:*
em_ph:*
em_pl:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    em_ro:*
em_sts:*
em_un:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                em_mu:*
em_om:*
em_or:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  gb_ba:*
gb_htg:*
gb_in:*
gb_om:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Gen Emb 1:*
                                                                                                                                        Perfect score:
Sequence:
                                                                                                                                                                            Scoring table:
                                                        OM protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Database :
                                                                                                                                                                                                                                                Searched:
                                                                                Run on:
```

```
Pestivirus type
                                      type
                                     Pestivirus
                                                                     Pestivirus
                                                                                                                                                              10864644
                                                                                                                                                                                                                                             source
             VERSION
KEYWORDS
SOURCE
ORGANISM
                                                                                                                                                                                               TITLE
JOURNAL
                                                                                                                                                 MEDLINE
PUBMED
                                                                                                                                                                        REFERENCE
AUTHORS
  ACCESSION
                                                                                REFERENCE
                                                                                             AUTHORS
                                                                                                                                      JOURNAL
                                                                                                                                                                                                                                FEATURES
                                                                                                                  TITLE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            VRL 12-JUL-2000
             PAT 20-APR-2002
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               652
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              532
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  100 AlaProGlnGlySerArgSerLeuThrProCysThrCysGlySerSerAspLeuTyrLeu 119
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             LeuSerProArgProlleSerTyrLeuLysGlySerSerGlyGlyProLeuLeuCysPro 159
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   833 CITTCGCCCCGGCCCATTTCCTACCTAAAAGGCTCCTCGGGGGGTCCGCTGTTGTGCCCC 892
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           160 AlaGlyHisAlaValGlyIlePheArgAlaAlaValCySThrArgGlyValAlaLySAla 179
                                                                                                                                                                                                                                                                                                                                                                                                                                                   583
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               120 ValThrArgHisAlaAspValIleProValArgArgArgGlyAspSerArgGlySerLeu 139
                                                                                                                                                                                                                                                                                                                                                                                 39
                                                                                                                                                                                                                                                                                                                                                                                                                             59
                                                                                                                                                                                                                                                                                                                                                                                                                                                                         79
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ProLysGlyProValIlcGlnMetTyrThrAsnValAspLysAspLeuValGlyTrpPro 99
                                                                                                                                                                                                                                                                                                                                                                                                                                                                        20 AlaGlnGlnThrArgGlyGluGluGlyCysGlnGiuThrSerGlnThrGlyArgAspLys
                                                                                                                                                                                                                                                                                                                                                                                                40 AsnGlnValGluGlyGluValGlnIleValSerThrAlaAlaGlnThrPheLeuAlaThr
                                                                                                                                                                                                                                                                                                                                                                                                                                         GlySerValValIleValGlyArgIleValLeuAsnGly-------AlaTyr
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Unknown.
Unclassified.
1 (bases 1 to 12734)
Hong,Z., Lai,V.C.H. and Lau,J.Y.N.
Hepatitis C virus protease-dependent chimeric pestivirus Patent: US 6326137-A 1 04-DEC-2001;
Location/Oualifiers
1.12734
Location/Oualifiers
1.12734
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ValAspPheIleProValGluSerLeuGluThrThrMetArgSer 194
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     linear
              linear
                                                                                                                                                                                                                                                     Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              RNA
                                                                                                                                                                                                                               Length:
Matches:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           AF268278 12734 bp
Pestivirus type 1, complete genome.
       Sequence 1 from patent US 6326137. AR179057 G1:20220612
                                                                                                                                                                                                                                                                                                              US-09-965-594-12 (1-195) x AR179057 (1-12734)
                                                                                                                                                                                                                                                                                        Gaps:
                                                                                                                                                                                                                               3.06e-56
928.50
94.87%
92.31%
90.94%
                                                                                                                                                                                                                                                    Percent Similarity:
Best Local Similarity:
                                                                     Unknown
                                                                                                                                                                                                                     Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             140
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      80
                                                                                                                                                                                                                                                                              Query Match:
                                             VERSION
KEYWORDS
SOURCE
ORGANISM
                                                                                                                                                               source
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      DEFINITION
                                                                                                                                                                                                                                Pred. No.:
                                                                                                               AUTHORS
TITLE
JOURNAL
                                                                                                                                                                                    BASE COUNT
                      DEFINITION
                                  ACCESSION
                                                                                                       REFERENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               AF268278
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       RESULT 2
                                                                                                                                                  FATURES
AR179057
                                                                                                                                                                                                ORIGIN
                                                                                                                                                                                                                                            Score:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              q
                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                          셤
                                                                                                                                                                                                                                                                                                                                                                                 ö
                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                 a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                             č
```

```
Direct Submission
Submitted (16-MAY-2000) Antiviral Therapy, Schering-Plough Research
Institute, 2015 Galloping Hill Road, Kenilworth, NJ 07033-0539, USA
Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         TEGIQRAMFQRGVNRSLHGIWPEKICTGVPSHLATDIELKTIHGMADASEKTNYTCCR
LQRHEWNKHGWCNWYNIEPWILVMNRTQANLTEGQPPRECAVTCRYDRASDLNVVTQA
RDSPTPLTGCKKGKNFSFAGILMRGPCNFEIAASDVLFKEHERISMFQDTTLYLVDGL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       KNCTPACLPKNTKIVGPGKFDTNAEDGKILHEMGGHLSEVLLLSLVVLSDFAPETASV
MYLLHFSIPQSHVDVADCDKTQLNLTVELTTADVIPGSVWNLGKWVCIRPNWMPYBT
TVVLAFEEVSOVVKLVLRALRDLTRINNAATTTAFLICLVKIVRGQWVGILWLLLT
GVQGHLDCKPESYRAIAKDERIGQLGABGLTTTWRSYSPGWKLDTWVIAWCEDGKLM
YLQRCTRETRYLALHTRALPRSVVFKKLFDGRKQEDVYEMNONFEFGLCPCDAKPIY
RGKFNTTLLNGPAFQWVCPIGWTGTVSCTSFNMDTLATTVVRTYRRSKPFPHRQGCIT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 QKNLGEDLHNCILGGNWTCVPGDQLLYKGGSIESCKWCGYQFKESEGLPHYPIGKCKL
ENETGYRLYDSTSCNREGVAIVPQGTLKCKIGKTTVQVIAMDTKLGPMPCRPYEIISS
EGPVEKTACTFNYTKTLKNKYFEPRDSYFQQYMLKGEYQYWFDLEVTDHHRDYFAESI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       LUVUVALLGGRYVLMLLUTYMVLSEORALGIOYGSGEVVMMGNLLTHNNIEVUTYFLL
LYLLIREESKKKWVLLLYHTLVVHPIRSYVILMIGDVYRADSGGGETLGRIDLET
LYVLLYIGLIIARRDPTIVPLYTIKANDIRSYVILMIGDVYRADSGGGETLGRIDLET
YFRYKKMLQCILSLVSGVFLIRSLIYLGRIEMPEVTIPNWRPLTLILLYLLSTTIVTR
WYDYAGLICQCVPILLLYTUMADELTILLLTTELTELYTTIVTRYNDSTYDSTYDOYDSGREYTLILLTTELTELTTIVTRYNGYNDSTYDDDSGEGEVYTLFRAFILISTTIVTR
IDYTRYDSIYDOYDGEGEGVYLFPSRORAGGGRISILLPLIRATISCYSSKWOLIYMSY
ILTLDFMYYMHRKVIEEISGGTNIISRIVAALIELNWSMEEEESKGLKKFYLLSGRIRN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ASSGRVVGRVKVGKNEESKPTKIMSGIQTVSKNTADLTEMVKKITSMNRGDFKQITLA
TGAGKTTELPKAVIEEIGRHKRVLVLIPLRAAAESVYQYMRLKHPSISFNLRIGDMKE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               LIIKHKVRNETVASHYGEEEVYGHPKIMTIIKASTLSKSRHCIICTVCEGREWKGGTC
PRCGRHGKPITCGMSLADFEERHYKRIFIREGNFFGMCSRCQGKHRRFEMDREPKSAR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    PCHISFGSRMPFROEYNGFVQYTARGOLFLRNLPVLATKVKMLMVGNLGEEIGNLEHL
GWILRGPAVCKKITEHEKCHINILDKLTAFFGIMPRGTTPRAPVRFPTSLLKVRRCLE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  TGWAYTHQGGISSVDHVTAGKDLLVCDSMGRTRVVCQSNNRLTDETEYGVKTDSGCPD
GARCYVLNPEAVNISGSKGAVVHLQKTGGEFTCVTASGTPAFFDLKNLKGWSGLPIFE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         LPDLDTV IDTGLKCEKRVRVSSK I PFIVTGLKRMAVTVGEQAQRRGRVGRVKPGRYYR
SQETATGSKDY HYDLLQAQRYG I EDG I NVTKSFREMNY DMSLY EEDSLLITQLEI LINN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   LLISEDLPAAVKNIMARTDHPEPIQLAYNSY EVOVPVLFPKIRNGEVTDTY ENYSELN
ARKLGEDVPVY IYATEDEDLAVDLLGLDWPDPGNOOVVETGKALKOVTGLSSAENALL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     TGRDKNQVEGEVQIVSTATQTFLATCINGVCWTVYHGAGTRTIASPKGPVIQMYTNVD
QDLVGWPAPQGSRSLTPCTCGSSDLYLVTRHANVIPVRRKGDSRGSLLSPRPISYLKG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SYSDTKEEGATKKKTQKPDRLERGKMKIVPKESEKDSKTKPPDATIVVEGVKYQVRKK
GKTKSKNTQDGLYHNKNKPQESRKKLEKALLAWAIIAIVLFQVTMGENITQWNLQDNG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    TNSLEGARQGTAKLTTWLGKQLGILGKKLENKSKTWFGAYAASPYCDVDRKIGYIWYT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   GDMATGITYASYGYFCQMPQPKLRAAWVEYSYIFLDEYHCATPEQLAIIGKIHRFSES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        /translation-"MELNTNEGSGSVVIVGRIVLSGSGSITACAQQTRGLLGCKITSL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SSGGPLLCPAGHAVGLFRAAVCTRGVAKAVDFIPVENLETTTRSGSGADTEDVVCCSM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       YCAECNRLHPAEEGDFWAESSMLGLKITYFALMDGKVYDITEWAGCQRVGISPDTHRV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                irvvamtatpagsvittgokhpieefiapevmkgedlgsofldiaglkipvdemkgnm
                                                                                                                                                                                                 no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  LVFVPTRNMAVEVAKKLKAKGYNSGYYYSGEDPANLRVVTSQSPYVIVATNAIESGV7
                                                                                                                                                                                                                                                                                1 (bases 1 to 12734)
Lai,V.C., Zhong,W., Skelton,A., Ingravallo,P., Vassilev,V.,
Donis,R.O., Hong,Z. and Lau,J.Y.
Generation and characterization of a hepatitis C virus NS3
protease-dependent bovine viral diarrhea virus
J. Virol. 74 (14), 6339-6347 (2000)
                                                                                                                                                                                            Viruses; ssRNA positive-strand viruses,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  /organism="Pestivirus type
/mol_type="genomic RNA"
/db_xref="taxon:11099"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              /product-"polyprotein"
/protein_id-"AAF82566.1"
/db_xref-"GI:9049957"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                /codon_start=1
AF268278.1 GI:9049956
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     2 (bases 1 to 12734)
Lai, V.C.H. and Hong, Z.
```

```
SerGlyGlyProLeuLeuCysProAlaGlyHisAlaValGlyIlePheArgAlaAlaVal 171
           AlaGlyHisAlaValGlyIlePheArgAlaAlaValCysThrArgGlyValAlaLySAla 179
                                                                                                                                                                                                                                                                   Length:
Matches:
                                                                                                                                                                                               21-NOV-2000;
                                                                                                                                                                                                                                 1145
                                                                                                AR118686 53 from patent US 6150087.
AR118686
                                                                                                                                                                                                                                                                                                                                                 LysLysGlySerValVallleValGly----
                                                                                                                                                                                                                                                                                                              Gaps:
                                                                                                                                                                                                                                                                                                                                US-09-965-594-12 (1-195) x AR118686 (1-5360)
                                                                                                                                                                    1 (bases 1 to 5360)
Chien, D.Y.
NANNY diagnostics and vaccines
Patent: US 6150087-A 53 21-NOV-
Location/Qualifiers
                                                                                                                                                                                                                       /organism="unknown"
a 1623 c 1532 g
                                                                                                                         AR118686.1 GI:14100596
                                                                                                                                                                                                                                                                 7.42e-64
894.50
89.71%
85.78%
87.61%
                                                                                                                                                                                                                 .5360
                                                                                                                                                             Unclassified.
                                                                                                                                                                                                                                                                                     Percent Similarity:
Best Local Similarity:
                                                                                                                                            Unknown.
                                                                                                                                                                                                                                 1060
                                                                                                                                                     Unknown
                                                                                                                                                                                                                                                           Alignment Scores:
                                                                                                                                                                                                                                                                                                                                               m
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 132
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1167
                                                                                                                                                                                                                                                                                                                                                                 867
                                                                                                                                                                                                                                                                                                                                                                                  18
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            92
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      152
           091
                                            180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1287
                                                                                                                                                                                                                                                                                                      Query Match:
                                                                                                                         VERSION
KEYWORDS
SOURCE
ORGANISM
                                                                                                                                                                                                                                                                  Pred. No.:
                                                                                                         DEFINITION
                                                                                                                                                                              AUTHORS
TITLE
JOURNAL
FEATURES
                                                                                                                                                                                                                                BASE COUNT
ORIGIN
                                                                                                                  ACCESSION
                                                                                                                                                                      REFERENCE
                                                                               RESULT 3
AR118686
                                                                                                                                                                                                                                                                            Score:
                                                                                                  LOCUS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              a
                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                 ó
                                                                                                                                                                                                                                                                                                                                                                                                   q
                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                       a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   đ
                       Q
                                            ò
                                                           a
                                                                                                                                                                                                                                                                                                                                                ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                οž
                                        VALEGYVGYQALSKRHVPMITDIYTIEDQRLEDTTHLQYAPNAIKTDGTETELKELAS
GDVEKIMGALSDYAAGGLEFWKSQARKIKTAPL-KENABAAKGYVOKFIDSLIENKEE
ITYGLWGTHTALYKSIAARLGHETAPATLUIKAAPGGESVSHYKCAAVDLVYYY
MNKDSPPGDSETOOGERRYVAELSGALATYTYKTWYTHVISKVVEPALAYLEYATSA
LKMFTPTRLESVVILSTTIYKTYLSIRKGKSDGILGTGISAAMFIISONPVSVGISVM
                                                                                                                                                                                                                                                                                                                                                                                           472
                                                                                                                                                                                                                                                                                                                                                                                                                                                              592
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 652
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   712
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      119
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      772
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ValThrargHisAlaAspValIleProValArgArgArgGlyAspSerArgGlySerLeu 139
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          832
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           LeuSerProArgProlleSerTyrLeuiysGlySerSerGlyGlyProLeuLeuCysPro 159
                                                                                                                                                                                                                                                                                                                                                                                                                             532
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            833 CTITCGCCCCGGCCCATTCCTACCTAAAAGGCTCCTCTGGGGGTCCGCTGTTGTGCCCC 892
                                                                                                                                                                                                                                                                                                                                                                                                           39
                                                                                                                                                                                                                                                                                                                                                                                                                                              59
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   66
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ProLysGlyProValIleGlnMetTyrThrAsnValAspLysAspLeuValGlyTrpPro
                                                                                                                                                                                                                                                                                                                                                                           -----AlaTyr
                                                                                                                                                                                                                                                                                                                                                                                  CCCAAGGGTCCTGTCATCCAGATGTATACCAATGTGGACCAAGACCTTGTGGGCTGGCCC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          GTTACGAGGCACGCCAACGTCATTCCCGTGCGCCGGCCAGGTGATAGCAGGGGTAGCCTG
                                                                                                                                                                                                                                                          1 others
                                                                                                                                                                                                                                                                                          12734
177
6
7
7
5
                                                                                                                                                                                                                                                                                                            Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                        GlySerValValIleValGlyArgIleValLeuAsnGly
                                                                                                                                                                                                                                                                                            Length:
Matches:
                                                                                                                                                                                                                                                                                                                                                       (1-12734)
                                                                                                                                                                                                                                                         3293 g
                                                                                                                                                                                                                                                                                                                                                       US-09-965-594-12 (1-195) x AF268278
                                                                                                                                                                                                                                                                                           1.35e-64
908.50
93.85%
90.77%
88.98%
                                                                                                                                                                                                                                                        2608 c
                                                                                                                                                                                                                                                  12509
                                                                                                                                                                                                                                                                                                          Percent Similarity:
Best Local Similarity:
Query Match:
                                                                                                                                                                                                                                                           ď
                                                                                                                                                                                                                                                          4030
                                                                                                                                                                                                                                                                                    Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                         S
                                                                                                                                                                                                                                                                                                                                                                                          413
                                                                                                                                                                                                                                                                                                                                                                                                                           473
                                                                                                                                                                                                                                                                                                                                                                                                                                             40
                                                                                                                                                                                                                                                                                                                                                                                                                                                              533
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    100
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      713
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         773
                                                                                                                                                                                                                                                                                                                                                                                                           50
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   80
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    653
                                                                                                                                                                                                                                                         BASE COUNT
ORIGIN
                                                                                                                                                                                                                                                  3'UTR
                                                                                                                                                                                                                                                                                             ٠.
و
                                                                                                                                                                                                                                                                                                    Score:
                                                                                                                                                                                                                                                                                                                                       DB
                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                         qq
                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                         d
                                                                                                                                                                                                                                                                                                                                                                                                                                            á
                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Op
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 d
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ογ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Š
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       셤
```

1047 GCTGCCCAAACCTTCCTGGCAACGTGCATCAATGGGGTGTGCTGGACTGTCTACCACGGG 1106 GACCAAGACCIIGIGGGCIGGCCCGCICCGCAAGGIAGCCGCICATIGACACCCIGCACI 1226 PAT 16-MAY-2001 -------AlaTyrAlaGlnGlnThrArgGlyGluGluGlyCysGlnGlu 31 52 AlaAlaGlnThrPheLeuAlaThrCysIleAsnGlyValCysTrpThrValTyrHisGly 71 :::::|||| ::::::|||| CGCAGGGCCGGGGGAGATGGTCTCCCAAGGGGTGGAGG AspLysAspLeuValGlyIrpProAlaProGlnGlySerArgSerLeuThrProCysThr ---ArglleValLeuAsnGly--ValaspPhelleProValGluSerLeuGluThrThrMetArgSer 194 linear 5360 175 8 10 11 Conservative: Mismatches: Indels: DNA ų

```
/organism="unknown"
1623 c 1533 g
                                                                                                                                                                                   1 (bases 1 to 5360)
Houghton, M., Choo, Q.-K.
Patent: WO 8904669-A 8 0
                                                                                                                                                                                                                                                                         7.42e-64
894.50
89.71%
85.78%
87.61%
                                                                      109328.1 GI:587963
                                                           192 MetArgSerPro 195
                                                                                                                                                                                                                      .5360
                                                                                                                                                                            Unclassified.
                                                                                                                                                                                                                                                                                                    Best Local Similarity:
                                                                                                                                                                                                                                        1061 a
                                                                                                                                                         Unknown
                                                                                                                                                                    Unknown
                                                                                                                                                                                                                                                                                          Percent Similarity:
                                                                                                                                                                                                                                                                  Alignment Scores:
Pred. No.:
                                                                                                                                                                                                                                                                                                                                                                                         18
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     112
                          172
                                                                                                                                                                                                                                                                                                          Query Match:
DB:
                                                                                                                                                                                                                        source
                                                                                                                         DEFINITION
                                                                                                                                                                  ORGANISM
                                                                                                                                                                                                                                        BASE COUNT
                                                                                                                                 ACCESSION
VERSION
                                                                                                                                                                                             AUTHORS
                                                                                                                                                                                     REFERENCE
                                                                                                                                                  KEYWORDS
                                                                                                                                                                                                               FEATURES
                                                                                               RESULT
I09328
                                                                                                                                                                                                                                                  ORIGIN
                                                                                                                                                          SOURCE
                                                                                                                                                                                                                                                                                   Score:
                                                                                                                LOCUS
                                                                                                                                                                                                                                                                                                                                                                       a
                                                                                                                                                                                                                                                                                                                                                                                                                                          5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                           q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ^{\circ}
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                         δ
                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                           õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ô
                          ò
                                           g
                                                           õ
                                                                             a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1346
                 GCCGGAACGAGGACCATCGCGTCACCCAAGGGTCCTGTCATCCAGATGTATACCAATGTA 1166
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  151
1347 TCGGGGGGTCCGCTGTTGTGCCCCGGGGGCACGCGCGTGGGCATATTTAGGGCCGCGGTG 1406
                                                                                                        PAT 02-DEC-1994
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SerGlyGlyProLeuLeuCysProAlaGlyHisAlaValGlyIlePheArgAlaAlaVal 171
                                                                                                                                                                                                                                                                                                                                                                                                          986
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                91
                                                                                                                                                                                                                                                                                                                                                      17
                                                                                                                                                                                                                                                                                                                                                                                         31
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                AlaGlyThrArgThrIleAlaSerProLySGlyProValIleGlnMetTyrThrAsnVal
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CGCAGGGGCCGGGAGATACTGCTCGGGCCAGCCGATGGAATGGTCTCCAAGGGGTGGAGG
                                                                                                                                                                                                                                                                                                                                                                                         132 ArgGlyAspSerArgGlySerLeuLeuSerProArgProIleSerTyrLeuLysGlySer
                                                                                                                                                                                                                                                                                                                                                      3 LysLysGlySerValVallleValGly-----ArgIleValLeuAsnGly---
                                                                                                         linear
                                                                                                                                                                                                                                                                         5360
175
8
10
11
                                                                                                                                                                                                                                                                         Length:
Matches:
Conservative:
Mismatches:
Indels:
Gaps:
                                                                                                         DNA
                                                                                                                                                                           1 (bases 1 to 5360)
Houghton, M., Choo, Q., f., and Kuo, G.
Nanby diagnostics and vaccines
Patent: EP 0318216-A1 48 31-MAY-1989;
                                                                                                       5360 bp
EP 0318216.
                                                                                                                                                                                                                                                                                                                                     US-09-965-594-12 (1-195) x 106434 (1-5360)
                                                                                                                                                                                                               Location/Qualifiers
1. .5360
                                                                                                                                                                                                                               /organism~"unknown"
i 1623 c 1533 g
                                                                                                               Sequence 48 from Patent
106434
106434.1 G1:590311
                                                                                                                                                                                                                                                                         7.42e-54
894.50
89.71%
85.78%
87.61%
                                                             1467 ATGAGGTCCCCG 1478
                                                    192 MetArgSerPro 195
                                                                                                                                                                   Unclassified
                                                                                                                                                                                                                                                                                                    Best Local Similarity:
                                                                                                                                                                                                                                         Ü
                                                                                                                                                   Unknown
                                                                                                                                                           Unknown
                                                                                                                                                                                                                                         1061
                                                                                                                                                                                                                                                                                           Percent Similarity:
                                                                                                                                                                                                                                                                  Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1107
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1167
                                                                                                                                                                                                                                                                                                                                                                                         18
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 72
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   92
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1287
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           152
                                                                                                                                                                                                                                                                                                        Query Match:
DB:
                                                                                                                                                                                                                         source
                                                                                                                 DEFINITION
                                                                                                                                                           ORGANISM
                                                                                                                                                                                                                                        BASE COUNT
                                                                                                                                                                                                                                                                           Pred. No.:
                                                                                                                        ACCESSION
VERSION
KEYWORDS
                                                                                                                                                                                    AUTHORS
TITLE
JOURNAL
                                                                                                                                                                             REFERENCE
                                                                                                                                                                                                               FEATURES
                                                                                                                                                                                                                                                  ORIGIN
                                                                                       RESULT
                                                                                                                                                   SOURCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                          q
                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Q
                                   g
                                                    ò
                                                                     đ
                                                                                                                                                                                                                                                                                                                                                     õ
                                                                                                                                                                                                                                                                                                                                                                                         õ
                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                   ò
```

```
1046
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        107 GCCGGAACGAGGACCATCGCGTCACCCAAGGGTCCTGTCATCCAGATGTATACCAATGTA 1166
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1407 TGCACCCGTGGAGTGGCTAAGGCGGTGGACTTTATCCCTGTGGAGAACCTAGAGACAACA 1466
                                                                                                                                                                                                                                  PAT 02-DEC-1994
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  986
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          131
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                151
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    152 SerGlyGlyProLeuLeuCysProAlaGlyHisAlaValGlyIlePheArgAlaAlaVal 171
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             17
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     51
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                91
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   31
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    32 ThrSerGlnThrGlyArgAspLysAsnGlnValGluGlyGluValGlnIleValSerThr
||||||||
987 ACCAGCCTAACTGGCCGGGACAAAAACCAAGTGGAGGGTGAGGTCCAGATTGTGTCAACT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              867 CGCAGGGGCCGGGAGATACTGCTCGGGCCAGCCGATGGAATGGTCTCCAAGGGGTGGAGG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ------AlaTyrAlaGlnGlnThrArgGlyGluGluGlyGSlnGlu
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    72 AlaGlyThrArgThrIleAlaSerProLysGlyProVallleGlnMetTyrThrAsnVal
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      92 AspLysAspLeuValGlyTrpProAlaProGlnGlySerArgSerLeuThrProCysThr
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                132 ArgGlyAspSerArgGlySerLeuLeuSerProArgProlleSerTyrLeuLysGlySer
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CysGlySerSerAspLeuTyrLeuValThrArgHisAlaAspValIleProValArgArg
                                                     CysThrArgGlyValAlaLysAlaValAspPheIleProValGluSerLeuGluThrThr
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ---ArgileValLeuAsnGly
                                                                                                                                                                                                                                  linear
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                5360
175
8
10
11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Conservative:
Mismatches:
                                                                                                                                                                                                                                  DNA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Length:
Matches:
                                                                                                                                                                                                                                                                                                                                                                                      M., Choo, Q.-K. and Kuo, G. to 8904669-A 8 01-JUN-1989; Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Indels:
Gaps:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1143
                                                                                                                                                                                                                         5360 bp
Sequence 8 from Patent WO 8904669.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             3 LysLysGlySerValValIleValGly--
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            US-09-965-594-12 (1-195) x I09328 (1-5360)
```

us-09-965-594-12.rge

02-DEC-1994

PAT

5

```
1140 CGCAGGGCCGGGAGATACTGCTCGGGCCAGCCGATGGAATGGTCTCCAAGGGGTGGAGG 1199
                                                                                                                                                                                                                                                                                                                                                                                                                                               ....------AlaTyrAlaGlnGlnThrArgGlyGluGluGlyCysGlnGlu 31
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        112 CysGlySerSerAspLeuTyrLeuValThrArgHisAlaAspValIleProValArgArg
ThrSerGlnThrGlyArgAspLysAsnGlnValGluGlyGluValGlnIleValSerThr
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   72 AlaGlyThrArgThrIleAlaSerProLysGlyProValIleGlnMetTyrThrAsnVal
                                                                                                                                                                                                                                                                                                                                                                                               3 LysLysGlySerValValIleValGly-----ArgIleValLeuAsnGly--
                                                                                                                                 linear
                                                                                                                                                                                                                                                                                                              6785
175
8
10
11
                                                                                                                                                                                                                                                                                                                               Conservative:
Mismatches:
Indels:
                                                                                                                                  DNA
                                                                                                                                                                                                         1 (bases 1 to 6785)
Houghton, M., Choo, Q.-L. and Kuo, G.
Nanby diagnostics and vaccines
Patent: EP 0318216-A1 54 31-MAY-1989;
                                                                                                                                                                                                                                                                                                               Length:
Matches:
                                                                                                                               106440 6785 bp 1
Sequence 54 from Patent EP 0318216.
106440 GI:590312
                                                                                                                                                                                                                                                                          1429
                                                                                                                                                                                                                                                                                                                                                                                US-09-965-594-12 (1-195) x IO6440 (1-6785)
                                                                                                                                                                                                                                            Location/Qualifiers
1. .6785
                                                                                                                                                                                                                                                                /organism="unknown"
2050 c 1914 g
                                                                                                                                                                                                                                                                                                              9.59e-64
894.50
89.71%
85.78%
87.61%
                                                                                   1740 ATGAGGTCCCCG 1751
                                                                          MetArgSerPro 195
                                                                                                                                                                                                  Unclassified.
                                                                                                                                                                                                                                                                                                                                 Percent Similarity:
Best Local Similarity:
                                                                                                                                                                                                                                                                          Ø
                                                                                                                                                                                                                                                                          1392
                                                                                                                                                                              Unknown
                                                                                                                                                                                         Unknown
                                                                                                                                                                                                                                                                                                        Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                      18
                                                                                                                                                                                                                                                                                                                                                                                                                                                                           32
                                                                           192
                                                                                                                                                                                                                                                                                                                                                     Query Match:
                                                                                                                                                                                                                                                         source
                                                                                                                                                   ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
                                                                                                                                LOCUS
DEFINITION
                                                                                                                                                                                                                                                                                                              Pred. No.:
                                                                                                                                                                                                                            TITLE
JOURNAL
                                                                                                                                                                                                                                                                          BASE COUNT
                                                                                                                                                                                                                                                                                                       Alignment
                                                                                                                                                                                                           REFERENCE
                                                                                                                                                                                                                    AUTHORS
                                                                                                              RESULT 7
IO6440
                                                                                                                                                                                                                                               FEATURES
                                                                                                                                                                                                                                                                                     ORIGIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            d
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       셤
                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               qc
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Ω
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                           g
                                                    q
                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1499
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ACCAGCCTAACTGGCCGGGACAAAACCAAGTGGAGGGTGAGGTCCAGATTGTGTCTAACT 1319
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                131
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 AspLysAspLeuValGlyTrpProAlaProGlnGlySerArgSerLeuThrProCysThr 111
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ArgGlyAspSerArgGlySerLeuLeuSerProArgProlleSerTyrLeuLysGlySer 151
                                                                                                                        PAT 16-MAY-2001
                             191
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            91
                                                                                                                                                                                                                                                                                                                                                                                                                             31
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  21
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      AlaAlaGlnThrPheLeuAlaThrCysIleAsnGlyValCysTrpThrValTyrHisGly 71
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            .......AlaTyrAlaGlnGlnThrArgGlyGluGluGlyCysGlnGlu
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CysGlySerSerAspLeuTyrLeuValThrArgHisAlaAspValIleProValArgArg
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             AlaGlyThrArgThrIleAlaSerProLysGlyProVallleGlnMetTyrThrAsnVal
ThrSerGlnThrGlyArgAspLysAsnGlnValGluGlyGluValGlnIleValSerThr
                             CysThrArgGlyValAlaLysAlaValAspPheIleProValGluSerLeuGluThrThr
                                                                                                                                                                                                                                                                                                                                                                                          ----ArglleValLeuAsnGly-
                                                                                                                         linear
                                                                                                                                                                                                                                                                                                     6785
175
8
10
11
                                                                                                                                                                                                                                                                                                                          Conservative:
                                                                                                                          DNA
                                                                                                                                                                                                                                                                                                                                 Mismatches:
Indels:
                                                                                                                                                                                            1 (bases 1 to 6785)
Chien,D.Y.
NANBV diagnostics and vaccines
Patent: US 6150087-A 65 21-NOV-2000;
Location/Qualifiers
1. 6785
                                                                                                                                                                                                                                                                                                       Length::
Matches:
                                                                                                                        6785 bp
US 6150087.
                                                                                                                                                                                                                                                                                                                                                                                          3 LysLysGlySerValVallleValGly---
                                                                                                                                                                                                                                                                                                                                                                        US-09-965-594-12 (1-195) x AR118692 (1-6785)
                                                                                                                                                                                                                                                    /organism-"unknown"
2050 c 1914 g
                                                                                                                        AR118692
Sequence 65 from patent
AR118692
                                                                                                                                                   AR118692.1 GI:14100602
                                                                                                                                                                                                                                                                                                      9.59c-64
894.50
89.71%
85.78%
87.61%
6
                                                                           MetArgSerPro 195
                                                                                                                                                                                         Unclassified
                                                                                                                                                                                                                                                                                                                         Percent Similarity:
Best Local Similarity:
                                                                                                                                                                                                                                                                  Ø
                                                                                                                                                                     Unknown.
Unknown.
                                                                                                                                                                                                                                                                                               Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1440
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1260
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1320
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1380
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  85
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        112
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1560
                                                                                                                                                                                                                                                                                                                                                                                                                              18
                                                                                                                                                                                                                                                                                                                                                                                                                                                                   32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        25
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             72
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             132
          1347
                                                                 192
                             172
                                                                                                                                                                                                                                                                                                                                           Query Match:
DB:
                                                                                                    RESULT 6
AR118692
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
                                                                                                                                                                                                                                                 source
                                                                                                                                                                     SOURCE
                                                                                                                                                                                                                                                                                                        ..
O
                                                                                                                                                                                                                                                                  BASE COUNT
                                                                                                                                                                                                                               JOURNAL
                                                                                                                                                                                                   REFERENCE
                                                                                                                                                                                                            AUTHORS
                                                                                                                                                                                                                                      PEATURES
                                                                                                                                                                                                                    TITLE
                                                                                                                                                                                                                                                                            ORIGIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       õ
                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                  qq
                                                                                                                                                                                                                                                                                                                                                                                         ò
                                             q
                                                                 ò
                             ò
```

1619

0y         152 SerGlyGlyProlecuLeuCysProAlaGlyHisAlaValGlyIlePheArgAlaAlaVal 171           0p         1620 TGGGGGGTCCGTGTTGGCCCGGGGGCACGCGGGGGTG 1679           0p         172 CysThrArgGlyValAlaLysAlaValAsPPheIleProValGluSerLeuGluThrThr 191           0p         172 CysThrArgGlyValAlaLysAlaValAsPPheIleProValGluSerLeuGluThrThr 191           1p         111111111111111111111111111111111111	Alignment Scores:     1.04e-63
	Secret Similarity: 89.59-64   Length: 6765     Secret Similarity: 89.713   Matches: 10     Best Local Similarity: 89.713   Conservative: 8     Best Local Similarity: 89.714   Conservative: 8     Best Local Similarity: 85.784   Mismatches: 10     Best Local Similarity: 87.784   Indeals: 11     Best Local Similarity: 87.784   Indeals: 12     Best Local Similarity: 87.784   Indeals: 12     Best Local Similarity: 87.784   Indeals: 13     Best Local Similarity: 87.784   Indeals: 13     Best Local Similarity: 87.784   Indeals: 13     Indeals: 11   Indeals: 12   Indeals: 13     Indeals: 12   Indeals: 13   Indeals: 13     Indeals: 13   Indeals: 13   Indeals: 13     Indeals: 13

Db 2085 CGGGGTGATAGCAGGGCAGCCTGCTGTCGCCCCGGCCCATTTCCTACTTGAAAGGCTCC 2144  Qy 152 SerGlyGlyProLeuLeuCysProAlaGlyHisAlaValGlyIlePheArgAlaAlaVal 171  [11] [11] [11] [11] [11] [11] [11] [1	HECPOLYP LOCUS HPCPOLYP LOCUS HPCPOLYP LOCUS DEFINITION Hepatitis C virus polyprotein gene, partial cds. ACCESSION M32084	ent 111	/ COUNTING STATE / / PEOTESING STATE / / AB_xref="GI:329876" / AB_	AUNTY CSCUMPALLIS SAEGALENI VILTENAS LAS IN ELVORITOR AND SAURAN	KADDE PPENELETTHRSPVETOUS SPPUYDS SCASKSTKVEADAGAGG  KAULD IND SYAATIGFGANTS ANGIDEN ITTGSPITYSTYGFELADGGGS  GGADI I TODECHSTDASIGGTCHOOPEN ANGARITHGSPITYSTYGFELADGGGS  GGADI I TODECHSTDASIGGTCHOOPEN ANGARIAGINAVATRATPRESUTVPHNIEE  VASSTTGET PFYCKAIPLEVIKGRHLIFCHSKKKODELAAKLVAGINAVATRGLD  VSYIPTSGDVVVVATDAIMTGYTGDFDSVIDGNTCUTOTVOTVDFLDPFITITIOD  AVSBTORRGREGREGRICH STRAYDAPGERPSGNEDSSVILGCYDAGCANYTEITPLOD  AVSBTORRGREGREGRICH STRAYDAPGERPSGNEDSSVILGCYDAGCANYTEITPLOD  AVSTWANTPGLPVODHLEFWEGVTTGLTHIDAHLSGYCHOSGCANYTEITPRAKHWMFIS  GHOVITAGGSTLAGGATGAATGAATGATGAATGAATGAATGAATGAATGA
Db         2085 CGGGGTGATAGCAGGGCAGCCTGCTGTCGCCCGGCCCATTTCCTACTTGAAAGGCTCC 2144           Qy         152 SerGlyGlyProLeuLcuCysProAlaGlyHisAlaValGLyILePhcArgAiaAlaVal 171           Db         2145 TGGGGGGTCGGTGTTGTGCCCGGGGGCAGCGCTGGGCATATTTAGGGCGGGTS 2204           Qy         172 CysThrArgGlyValAlaLysAlaValAspPheTleProValGluSerLeuGluThrThr 191           Db         1205 TGCACCGTGGAGTGGTAAGGCGGTGGACTTTATCCCTGTGGAGAACCTAGAGACAACC 2264           Qy         192 McArgGsrPro 195           IIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIII	No.	Alignment Scores: 1.04c-63 Longth: 7310  Pred. No.: 894.50 Matches: 175  Score: 894.50 Matches: 175  Percent Similarity: 89.71% Conservative: 8  Best Local Similarity: 85.78% Mismatches: 10  Query Match: 6 Indels: 2  US-09-965-594-12 (1-195) x 109331 (1-7310)	Qy         3 LysLysGlySerValValIleValGly	Qy         3.2 ThrSerGlnThrGlyArgAspLysAsnGlnValGluGlyGiuValGlnIleValScrThr 51           Db         1785 ACCAGCCTAACTGGCGGGACAAAACCAAGTGGAGGTGAGGTCCAGATTGTGTCAACT 1844           Qy         5.2 AlaAlaGlnThrPheLeulaThrCysIleAsnGlyValCysTrpThrValTyrHisGly 71           Db         1845 GCTGCCCAAACCTTGCTGGAACGTGCATCATGGGGTGTGCTGGACTGTTACCACGG 1904	Oy 72 AlaGlyThrArgThrIleAlaSerProLysGlyProVallleGlnMetTyrThrAsnVal 91

```
source
AR118703
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
                                                                                                                                                                                               BASE COUNT
ORIGIN
                                                                                    ORGANISM
                                                                                                           REFERENCE
AUTHORS
TITLE
JOURNAL
                                                                                                                                                               FEATURES
                                                                          SOURCE
                                                                                                                                                                                                                                                             Score:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                d
                                                                                                                                                                                                                                                                                                                                                                                                           òγ
                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                         ^{\circ}
                                                                                                                                                                                                                                                                                                                                                                                                                                                                              q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   οy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ô
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
          PEFFELDGVRLHRPAPPCKPLLREEVSFRVGLHEYPVGSQLPCEPEPDVAVLTSMLT
DPSHITABAAGKRAARSPSVASSSASQLSAPSLAATCTAHNDSPDAELIBANLJWR
QEWGGNITRVESENKVYILDSFDPLVAEEDEREISVPAETLRKSRRFAQALPVWARPD
YNPPLVETWKRPDYEBPWYWGCPLPPPKSPPYPPPRKKRTVYJTESTLSTALAELATR
SPGSSTGSITGDNTTTSSEPAPSGCPPDSDAESYSMPPLGFEDDPLLSOSUSTV
SSENARDVVCCAMSYSWIGALYPCAAESCKLPINALSNELLRHINLVYSTTSRSG
QRQKKYTPORLQVLDSHYODVLKEVKAAASKVKANLLSVEEACSLTPPHSAKSKFGYG
AKDVRCHARKAYTHINSVWXDLLEDNYTPIDTIMAKNEVFCVOPEKGGRRPRLIVF
PDLGVYRCEKMALTVOVTKLPLEAMGSYGFOYSFGOYSFGOYPECGSTRPRLIVF
RCPDSTYTESDIRFEEATVOCCDLDPQARVAIKSLTWGGPLTMSGENGGYRR
CRASGVLTTSCGNTLTCYIKARAACRAGLODCTMLVCGDDLYVICESAGVQEDAASL
RAFTERATRYSAPPODPQPRYDLELITSCSSNVSVAHDGAGKRYYLTRDPTPLAR
AARTARHTPVNSULGNIIMFAPTLMARMILMTHFFSVLIARDQLEGYGGCY
SIEPLDLPPIIQRL

SIEPLDLPPIIQRL

SIEPLDLPPIIQRL

SIEPLDLPPIIQRL

SIEPLDLPPIIQRL

SIEPLDLPPIIQRL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            2024
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    2084
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    2144
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         2205 TGCACCCGTGGAGTGGCTAAGGCGGTGGACTTTATCCCTGTGGAGAACCTAGAGACAACC 2264
                                                                                                                                                                                                                                                                                                                                                                     1665 CGCAGGGGCCGGGAGATACTGCTCGGGCCAGCCGATGGAATGGTCTCCAAGGGGTGGAGG 1724
                                                                                                                                                                                                                                                                                                                                                                                                           ACCAGCCTAACTGGCCGGGACAAAACCAAGTGGAGGGTGAGGTCCAGATTGTGTCAACT 1844
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             151
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             191
  PINAYTTGPCTPLPAPNYTFALWRVSAEEYVEIRQVGDFHYVTGMTTDNLKCPCQVPS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               92 AspLysAspLeuValGlyTrpProAlaProGlnGlySerArgSerLeuThrProCysThr 111
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CysGlySerSerAspLeuTyrLeuValThrArgHisAlaAspVallLeProValArgArg 131
                                                                                                                                                                                                                                                                                                                                                                                             21
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             71
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ArgGlyAspSerArgGlySerLeuLeuSerProArgPro1leSerTyrLeuLysGlySer
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             AlaAlaGlnThrPheLeuAlaThrCysIleAsnGlyValCysTrpThrValTyrHisGly
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           2085 CGGGGTGATAGCAGGGGCAGCCTGCTGTCGCCCCGGCCCATTTCCTACTTGAAAGGCTCC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            LysLysGlyServalVallleValGly-----ArglleValLeuAsnGly-----
                                                                                                                                                                                                                                                                                                                                                                                                                                              ThrserGlnThrGlyArgAspLysAsnGlnValGluGlyGluValGlnIleValSerThr
                                                                                                                                                                                                                                   7310
175
8
10
11
                                                                                                                                                                                                                                   Length:
Matches:
Conservative:
Mismatches:
Indels:
Gaps:
                                                                                                                                                                                     2058 9 1539
                                                                                                                                                                                                                                                                                                                        US-09-965-594-12 (1-195) x HPCPOLYP (1-7310)
                                                                                                                                                                                                                                  1.04e-63
894.50
89.71%
85.78%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     2265 ATGAGGTCCCCG 2276
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               MetArgSerPro 195
                                                                                                                                                                                      2218 c
                                                                                                                                                                                                                                                     Percent Similarity:
Best Local Similarity:
Query Match:
DB:
                                                                                                                                                                                      1495
                                                                                                                                                                                                                         Alignment Scores:
Pred. No.:
                                                                                                                                                                                                                                                                                                                                               ٣
                                                                                                                                                                                                                                                                                                                                                                                                                                               32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1785
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             52
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1845
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               112
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             132
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             172
                                                                                                                                                                                                                                                                                                                                                                                                18
                                                                                                                                                                                     BASE COUNT
ORIGIN
                                                                                                                                                                                                                                                  Score:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ογ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                              ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                a
                                                                                                                                                                                                                                                                                                                                                                                                                                         δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  QQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ογ
```

```
2730
                                                                                                                                                                                                                                                                                                                                                                                                                                                        2790
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         2850
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              2910
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  2970
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    3030
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       3090
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           3150
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              3210
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      PAT 16-MAY-2001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           111
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               131
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 151
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     171
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         191
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  21
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    7
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        91
                                                                                                                                                                                                                                                                                                                                                                                                                              31
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       2671 CGCAGGGGCCGGGAGATACTGCTCGGGCCAGGCCGATGGAATGGTCTCCAAGGGGTGGAGG
                                                                                                                                                                                                                                                                                                                                                                                                                                          2731 TIGCIGGCCCCATCACGCCGTACGCCCAGCAGAGAGGGCCTCCTAGGGTGCATAAIC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1151 TCGGGGGGTCCGCTGTTGTGCCCCGCGGGGCACGCCGTGGGCATATTAGGGCCGCGGTG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         172 CysThrArgGlyValAlaLysAlaValAspPheIleProValGluSerLeuGluThrThr
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              AspLysAspLeuValGlyTrpProAlaProGlnGlySerArgSerLeuThrProCysThr
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         2971 GACCAAGACCTTGTGGGCTGGCCCGCTCCGCAAGGTAGCCGCTCATTGACACCTGCTGCGCTC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     152 SerGlyGlyProLeuLeuCysProAlaGlyHisAlaValGlyIlePheArgAlaAlaVal
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     AlaalaGlnThrPheLeuAlaThrCysIleAsnGlyValCysTrpThrValTyrHisGly
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ArgGlyAspSerArgGlySerLeuLeuSerProArgProlleSerTyrLeuLysGlySer
                                                                                                                                                                                                                                                                                                                                                                                                                              -------AlaTyrAlaGlnGlnThrArgGlyGluGluGlyCysGlnGlu
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ThrSerGlnThrGlyArgAspLysAsnGlnValGluGlyGluValGlnIleValSerThr
                                                                                                                                                                                                                                                                                                                                                                            ---ArgIleValLeuAsnGly
      linear
                                                                                                                                                                                                                                                      8316
175
8
8
10
11
                                                                                                                                                                                                                                                                    Matches:
Conservative:
Mismatches:
Indels:
      DNA
                                                                                                         (dases 1 to 8316)
Chien, D.Y.
NANBV diagnostics and vaccines
Patent: US 6150087-A 88 21-NOV-2000;
Location/Qualifiers
8316 bp 1
Sequence 88 from patent US 6150087.
                                                                                                                                                                                                                                                          Length:
                                                                                                                                                                                                                                                                                                                                                                           LysLysGlySerValValIleValGly----
                                                                                                                                                                                                                                                                                                                         Gaps:
                                                                                                                                                                                                                                                                                                                                                  (1-8316)
                                                                                                                                                                                        /organism="unknown"
2529 c 2345 g
                                                                                                                                                                                                                                                                                                                                                 US-09-965-594-12 (1-195) x AR118703
                                            AR118703.1 GI:14100613
                                                                                                                                                                                                                                                       1.2e-63
894.50
89.71%
85.78%
87.61%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           MetArqSerPro 195
                                                                                                                                                                             .8316
                                                                                                Unclassified.
                                                                                                                                                                                                                                                                             Percent Similarity:
Best Local Similarity:
Query Match:
DB:
                                                                                                                                                                                                      1671 a
                                                                .
Unknown.
                                                                                     Unknown
                                                                                                                                                                                                                                           Alignment Scores:
Pred. No.:
                                                                                                                                                                                                                                                                                                                                                                           ٣
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     22
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           92
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 132
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              192
                                                                                                                                                                                                                                                                                                                                                                                                                              18
```

```
3811
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                3391
                                                                                                                                                                                                                                                                                                                                                                 3571
                                                                                                                                                                                                                                                                                                                                                                                                3631
                                                                                                                                                                                                                                                                                                                                                                                                                                 3691
                                                                                                                                                                                                                                                                                                                                                                                                                                                3451
                                                                                                                                                                                                                                                                                                                                                                                                                111
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  191
                PAT 16-MAY-2001
                                                                                                                                                                                                                                                                                                                                                7.1
                                                                                                                                                                                                                                                                                                                                                                                91
                                                                                                                                                                                                                                                                                 31
                                                                                                                                                                                                                                                17
                                                                                                                                                                                                                                                                               52 AlaAlaGlnThrPheLeuAlaThrCysIleAsnGlyValCysTrpThrValTyrHisGly
                                                                                                                                                                                                                                                                                                                                                          CysThrArgGlyValAlaLysAlaValAspPheIleProValGluSerLeuGluThrThr
                                                                                                                                                                                                                                                                 332 CGCAGGGGCCGGGAGATAÇTGCTCGGGCCAGCCGATGGAATGGTCTCCAAGGGGTGGAGG
                                                                                                                                                                                                                                                LysLysGlySerValValIleValGly-----ArgIleValLeuAsnGly-----
                linear
                                                                                                                                                                        9185
175
8
10
11
                                                                                                                                                                        Length:
Matches:
Conservative:
Mismatches:
Indels:
                                                                              1 (bases 1 to 9185)
Chien, D.Y.
NANBV diagnostics and vaccines
Patent: US 6150087-A 122 21-NOV-2000;
Location/Qualifiers
1. 9185
               9185 bp E
from patent US 6150087.
                                                                                                                                                                                                                                (1-9185)
                                                                                                                                 /organism="unknown"
2790 c 2608 g
                                                                                                                                                                                                                                 US-09-965-594-12 (1-195) x AR118722
                                AR118722
AR118722.1 GI:14100632
                                                                                                                                                                         1.33e-63
894.50
89.71%
85.78%
87.61%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    MetArgSerPro 195
                    Sequence 122
AR118722
                                                                         Unclassified.
                                                                                                                                                                                          Percent Similarity:
Best Local Similarity:
                                                                                                                                          ಗ
                 AR118722
                                                                                                                                         1849
                                                         Unknown
                                                                 Unknown
                                                                                                                                                                 Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    192
                                                                                                                                                                                                          Ouery Match:
RESULT 14
AR118722
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
                                                        SOURCE
ORGANISM
                                                                                                                                         BASE COUNT
ORIGIN
                                                                                                                                                                           Pred. No.:
                                                                                        AUTHORS
TITLE
JOURNAL
                                                                                 REFERENCE
                                                                                                                 FEATURES
                                                                                                                                                                                    Score:
                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                               qq
                                                                                                                                                                                                                                                                                                                                                                                                                 δλ
                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                 qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                 ö
                                                                                                                                                                                                                                                                                                qq
                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                 á
                                                                                                                                                                                                                                                                                                                 9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                          3072
                                                                                                                                                                                                                                                                                                                                                                          AspLysAspLeuValGlyTrpProAlaProGlnGlySerArgScrLeuThrProCysThr 111
                                                                                                                                                                                                                                                                                                                                                                                                                                          CysGlySerSerAspLeuTyrLeuValThrArgHisAlaAspValIleProValArgArg 131
         PAT 16-MAY-2001
                                                                                                                                                                                                                                                                                                                                          71
                                                                                                                                                                                                                                                                         31
                                                                                                                                                                                                                                                                                                         51
                                                                                                                                                                                                                                                    52 AlaAlaGlnThrPheLeuAlaThrCysIleAsnGlyValCysTrpThrValTyrHisGly
                                                                                                                                                                                                                                                                          -------AlaTyrAlaGlnGlnThrArgGlyGluGluGlyCysGlnGlu
                                                                                                                                                                                                                                                                                   ThrSerGlnThrGlyArgAspLysAsnGlnValGluGlyGluValGlnIieValSerThr
                                                                                                                                                                                                                                         LysLysGlySerValValIleValGly-----ArgIleValLeuAsnGly--
          linear
                                                                                                                                                                 8987
175
8
10
11
                                                                                                                                                                 Length:
Matches:
Conservative:
Mismatches:
Indels:
                                                               Unclassified.
1 (bases 1 to 8987)
Chien,D.Y.
NAMBV diagnostics and vaccines
Patent: US 6150087-A 137 21-NOV-2000;
         AR118728 8987 bp 5
Sequence 137 from patent US 6150087.
AR118728.1 GI:14100638
                                                                                                                                                                                                           Gaps:
                                                                                                                                                                                                                           US-09-965-594-12 (1-195) x AR118728 (1-8987)
                                                                                                                        /organism≂"unknown"
2735 c 2547 q
                                                                                                                                                                 1.3e-63
894.50
89.71%
85.78%
87.61%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      MetArgSerPro 195
                                                                                                                    .8987
                                                                                                                                                                                  Percent Similarity:
Best Local Similarity:
                                                                                                                                    ď
                                                  Unknown.
                                                           Unknown
                                                                                                                                                           Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           3433
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            192
                                                                                                                                                                                                                                                                                                                                                                                                                          3313
                                                                                                                                                                                                                                                                                                                          3133
                                                                                                                                                                                                                                                                                                                                                           3193
                                                                                                                                                                                                                                                                                                                                                                           72
                                                                                                                                                                                                                                                                                                                                                                                                                                           112
                                                                                                                                                                                                                                           ٣
                                                                                                                                                                                                                                                                          18
                                                                                                                                                                                                                                                                                                          32
                                                                                                                                                                                                                                                                                                                                                                                           3253
                                                                                                                                                                                                                                                                                                                                                                                                           92
                                                                                                                                                                                                  Query Match:
DB:
                         ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
                                                                                                                    source
          LOCUS
                                                                                                                                  BASE COUNT
                                                                                                                                                                    ..
9
                                                                                 AUTHORS
TITLE
JOURNAL
                                                                          REFERENCE
AR118728
                                                                                                          FEATURES
                                                                                                                                           ORIGIN
                                                                                                                                                                    Pred.
                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                        Op
                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           q
```

QQ ò

g

ö

qq

οχ

q ò

ò

ò

ò

```
8
                                                                                                                                                                                                                                                                                                                                                                                                       5495
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 5794 TIGCIGGCGCCCATCACGGCGTACGCCCAGCAGAGGGGCCTCCTAGGGGGGCATAAIC 5735
                                                                                                                                                                                                                                                                                                                                                                            5854 CGCAGGGGCCGGGAGATACTGCTCGGGCCAGCCGATGGAATGGTCTCCAAGGGGTGGAGG 5795
                                                                                                                                                                                                                                                                                                                                                                                                                                                                              AspLysAspLeuValGlyTrpProAlaProGlnGlySerArgSerLeuThrProCysThr 111
                                    PAT 16-MAY-2001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                172 CysThrArgGlyValAlaLysAlaValAspPheIleProValGluSerLeuGluThrThr 191
                                                                                                                                                                                                                                                                                                                                                                                                                                            AlaGlyThrArgThrIleAlaSerProLySGlyProVallleGlnMctTyrThrAsnVal 91
                                                                                                                                                                                                                                                                                                                                                                   32 ThrSerGlnThrGlyArgAspLysAsnGlnValGluGlyGluValGlnIleValSerThr 51
                                                                                                                                                                                                                                                                                             3 LysLysGlySerValValIleValGly-----17
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ------AlaTyrAlaGlnGlnThrArgGlyGluGluGlyCysGlnGlu
                                     linear
                                                                                                                                                                                                           9185
175
10
10
11
                                                                                                                                                                                                           Length:
Matches:
Conservative:
Mismatches:
                                                                                                          Chien, D.Y.

NANBV diagnostics and vaccines
Patent: US 6150087-A 123 21-NOV-2000;
1. .9185
                                   AR118723 9185 bp D
Sequence 123 from patent US 6150087.
AR118723.1 G1:14100633
                                                                                                                                                                                                                                                         Gaps:
                                                                                                                                                                                                                                                                           US-09-965-594-12 (1-195) x AR118723 (1-9185)
                                                                                                                                                               /organism="unknown"
a 2608 c 2790 g
                                                                                                                                                                                                            1.33e-63
894.50
89.71%
85.78%
87.61%
3932 ATGAGGTCCCCG 3943
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     MetArgSerPro 195
                                                                                                    Unclassified.
                                                                                                                                                                                                                            Percent Similarity:
Best Local Similarity:
Query Match:
DB:
                                                                             Unknown.
                                                                                          Unknown
                                                                                                                                                                                                    Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                             5614
                                                                                                                                                                                                                                                                                                                                                                                                                                           72
                                                                                                                                                                                                                                                                                                                                                                                                                                                                              92
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     192
                                                                                                                                                                                                                                                                                                                                18
                                                                                                                                                         source
                                            DEFINITION
ACCESSION
VERSION
KEYWORDS
                  RESULT 15
AR118723/c
LOCUS
                                                                                                                                                                                                             No.:
                                                                                          ORGANISM
                                                                                                                                                                        BASE COUNT
                                                                                                          REFERENCE
AUTHORS
TITLE
                                                                                                                                     JOURNAL
                                                                                                                                                                                  ORIGIN
                                                                                                                                                                                                                      Score:
                                                                                                                                                                                                                                                                                             ŏ
                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                a
                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                         D
                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      οζ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                ôy
```

```
1111111111
5254 ATGAGGTCCCCG 5243
```

Search completed: August 31, 2003, 00:45:40 Job time : 2558.57 secs

```
GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.
```

```
(without alignments)
2905.924 Million cell updates/sec
                                                                                                                                                       MKKKGSVVIVGRIVLNGAYA......VAKAVDFIPVESLETTMRSP 195
                                      ; Search time 181.082 Seconds
OM protein - nucleic search, using frame_plus_p2n model
                                    August 30, 2003, 19:13:57
                                                                                                      US-09-965-594-12
1021
                                                                                                              Title:
Perfect score:
Sequence:
                                                                                                                                                                                                Scoring table:
                                      Run on:
```

2552756 seqs, 1349719017 residues 0.5 Xgapext Ygapext BLOSUM62 Xgapob 10.0 , Xgapext Ygapop 10.0 , Ygapext Fgapop 6.0 , Fgapext Delop 6.0 , Delext , Fgapext Searched:

5105512 Total number of hits satisfying chosen parameters:

Post-processing: Minimum Match 0% Maximum Match 100% Listing first 45 summaries length: 0 length: 2000000000 Minimum DB seq Maximum DB seq

Command line parameters:

-WODEL-Itame+\_pro.model - DEV-xlp
'Q-Cap21\_VICSPTO\_RODO\_IVES0965594/runat\_29082003\_151918\_28302/app\_query.fasta\_1.2872
'Q-Cap22\_IVESPTO\_RODO\_IVES09655594/runat\_29082003\_151918\_28302/app\_query.fasta\_1.2872
'LOOPEXT-0. UNITS-bits - START-1 - END--1 - MATHIX-blosum62 - TRANS-human40.cdi
-LIST-45 - DOCALIGN-200 - THR\_SCORE-POCT - THR\_MARX-100 - THR\_MIN-0 - ALIGN-15
-USER-USO9965594\_eCGN\_1 1.1412\_@runat\_29082003\_151918\_28302 - NCPU-6 - ICPU-3
-NO\_MARP - LARGEQUERY - NEG. SCORES-0 - WALT - DSPBLCCK-100 - LONGLOG
- FRANCOUR-120 - WARN - TAMEOUT-30 - THREADS-1 - XCAPOP-10 - XCAPEXT=0.5 - FGAPOP-6
- FGAPEXT-7 - YCAPOP-10 - YGAPEXT-0.5 - DELEXT-7

Database

```
N_Geneseq_19Jun03:*

1. $IDSI,gcgdata/geneseq.geneseqn.embl/NA1980.DAT:*

2. $IDSI,gcgdata/geneseq.geneseqn.embl/NA1981.DAT:*

4. $IDSI,gcgdata/geneseq.geneseqn.embl/NA1981.DAT:*

5. $IDSI,gcgdata/geneseq.geneseqn.embl/NA1981.DAT:*

6. $IDSI,gcgdata/geneseq.geneseqn.embl/NA1981.DAT:*

7. $IDSI,gcgdata/geneseq.geneseqn.embl/NA1981.DAT:*

7. $IDSI,gcgdata/geneseq.geneseqn.embl/NA1980.DAT:*

8. $IDSI,gcgdata/geneseq.geneseqn.embl/NA1980.DAT:*

9. $IDSI,gcgdata/geneseq.geneseqn.embl/NA1980.DAT:*

10. $IDSI,gcgdata/geneseq.geneseqn.embl/NA1980.DAT:*

11. $IDSI,gcgdata/geneseq.geneseqn.embl/NA1991.DAT:*

12. $IDSI,gcgdata/geneseq.geneseqn.embl/NA1991.DAT:*

13. $IDSI,gcgdata/geneseq.geneseqn.embl/NA1991.DAT:*

14. $IDSI,gcgdata/geneseq.geneseqn.embl/NA1991.DAT:*

15. $IDSI,gcgdata/geneseq.geneseqn.embl/NA1991.DAT:*

16. $IDSI,gcgdata/geneseq.geneseqn.embl/NA1991.DAT:*

17. $IDSI,gcgdata/geneseq.geneseqn.embl/NA1991.DAT:*

18. $IDSI,gcgdata/geneseq.geneseqn.embl/NA1991.DAT:*

20. $IDSI,gcgdata/geneseq.geneseqn.embl/NA1999.DAT:*

21. $IDSI,gcgdata/geneseq.geneseqn.embl/NA1999.DAT:*

22. $IDSI,gcgdata/geneseq.geneseqn.embl/NA1999.DAT:*

23. $IDSI,gcgdata/geneseq.geneseqn.embl/NA1999.DAT:*

24. $IDSI,gcgdata/geneseq.geneseqn.embl/NA1999.DAT:*

25. $IDSI,gcgdata/geneseq.geneseqn.embl/NA1999.DAT:*

26. $IDSI,gcgdata/geneseq.geneseqn.embl/NA1999.DAT:*

27. $IDSI,gcgdata/geneseq.geneseqn.embl/NA1999.DAT:*

28. $IDSI,gcgdata/geneseq.geneseqn.embl/NA1999.DAT:*

29. $IDSI,gcgdata/geneseq.geneseqn.embl/NA1999.DAT:*

20. $IDSI,gcgdata/geneseq.geneseqn.embl/NA1999.DAT:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2001B.DAT:*/SIDS1/gcgdata/geneseq/geneseqn-embl/NA2002.DAT:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    /SIDS1/gcgdata/geneseq/genesegn-emb1/NA2001A.DAT:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         /SIDS1/gcgdata/geneseg/genesegn-embl/NA2003.DAT:*
```

RESULT 1

AAA73329

No. is the number of results predicted by chance to have a greater than or equal to the score of the result being printed, Pred.

is derived by analysis of the total score distribution.

SUMMARIES

		æ				
Result		Query				
. ON .	Score	Match	Length	DB	ID	ription
	100	100	005		, ,	
4 6	<b>V</b> (	3 5	0 (	7 7	700	ILIS C VIFU
7 (	<b>π</b> (	٠,	יה הו	77	2 1 2	itis C
η.	D (	۰ ۵	ñ	77	332	itis C
4	oo -	9	65	21	333	itis
'n	ഗ	m		21	333	
9	ഗ	m	29	2,	3	itis
7	ന	C4	93	21	m	ပ
ω	8	-	59	21	33	is
6	28.	0	1273	24	ü	BVDV/HC
10	18.	90.0	61	25	ABX15706	al
11	94.	~	530	10	AAN92097	ined open
12	94.	~	536	10	AAN90327	,0
13	94.	7	069	10	AAN92103	open
14	94.	~	731	10	AAN92106	open
15	94.	$\sim$	731	10	AAN90336	hepat
16	94.	~	731	16	AA098221	Lis
17	94	~	831	21	AAA75296	ĝ
18	94.	~	913	50	AA207656	Nucleotide sequenc
19	94	~	918	11	AA005956	e stra
20	94.	~	918	12	AA010566	Litis C vir
21	94.	~	918	21	AAA75297	. 0
22	94.	^	940	13		led HCV CD
23	94.	7.	940	17	AAT12710	itis C v
24	•	7	940	18	AAT99981	prot
25	94.	7	940	19		rotein
26	94.	ζ.	940	24		Hepatitis C virus
27	σ.	ζ.	20	21	AAA70344	
28	894	87.6	2058	24	5	3/4
29	മ	ζ.	205	24	534	Hepatitis C virus
30	8	٠,	205	52	ABX14410	ğ
3	•	ς.	950	15	AAQ74770	is C vi
32	ഗ	۲.	193	20	AAX23258	V NS3 DNA.
(5)	8	ζ.	874	20	AAX23259	PET-BS
34		ζ.	189	20	AAX80355	-NS3
35	91.	ζ.	918	50	AAX26737	Nucleotide sequenc
36	91.		918	20	AAX00459	itis C
3.	90.	ζ.	831	11	95	tis C
88	ത	۲.	206	24		itis
39	8	'n	206	24	AAD31767	Hepatitis C virus
0	89.	۲.	964	19		citis C
	89	ζ.	964	24	28	CDNA encoding hepa
	68	7	1298	13	2936	atitis C vi
		87.1	1298	77	K8728	titi
4 4	89.	ζ.	1662	21	3621	ednen
₹.	88.	7	199	20	2	HCV NS4A-NS3 compl

Hepatitis C virus NS4A-NS3 fusion protease coding sequence #2. ALIGNMENTS BP AAA73329 standard; DNA; 588 (first entry) 19-DEC-2000 AAA73329;

Hepatitis; NS3 protease; viral replication; chronic liver disease; liver failure; liver cancer; mutant; mutein; ds. Hepatitis C virus. Synthetic.

Location/Qualifiers

Key

a

```
CCGCAGGGTTCCCGTTCCCTGACCCCGTGCACCTGCGGTTCCTCCGACCTGTACCTGGTT 360
                                                                                                                                                                                                                                                                                                                                                                                                                                  13-JUL-2000
                                                                                                                                                181
                                                                                                       161
                                                                                                                                                                                                                                 AAA73330
                                                                                                                                                                                       RESULT 2
                                                                                                                                                                                                  AAA73330
                                                                                                                                                                                                              οq
                                                              ô
                                                                                   a
                                                                                                        ò
                                                                                                                          g
                                                                                                                                                ò
                                                                                                                                                                  g
                      õ
                                                                                                                                                                                                                                                                       The present sequence is the coding sequence for a mutated version of a fusion protein created using the Hepatitis C virus (HCV) NS3 and NS4A protease enzymes. These proteins are both essential for the replication of the virus, acting to cleave its replicative proteins from the POP proteins from the POP genome. Inhibitors of the two proteins should be effective as antiviral treatments of HCV infection. This is useful as HCV can load to chronic liver disease such as cirrhosis, liver failure and liver cancer. The present invention concerns a number of NS3 mutants and NS3-NS4A fusion proteins which can be used to identify inhibitors of this type, as well as enabling structural studies of the protease and protease: inhibitor complexes. The protein produced from this
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ProGlnGlySerArgSerLeuThrProCysThrCysGlySerSerAspLeuTyrLeuVal 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     61 CAGCAGACTCGAGGTGAGGAGGGTTGCCAAGAACCTCCCAGACCGGTCGTGACAAAAAC 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            LysGlyProVallleGlnMetTyrThrAsnValAspLysAspLeuValGlyTrpProAla 100
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         09
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 80
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                40
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             9
                                                                                                                                                                                                 Modified hepatitis C virus (HCV) NS3 protease comprising at least l substitution of a hydrophobic alpha-helix 0 amino acid to a hydrophilic amino acid, useful for screening inhibitors that may treat hepatitis \mathbb C
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       GlnValGluGlyGluValGlnIleValSerThrAlaAlaGinThrPheLeuAlaThrCys
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  61 IleAsnGlyValCysTrpThrValTyrHisGlyAlaGlyThrArgThrIleAlaSerPro
1..588
/*tag= a
/product= "NS4A-NS3 fusion protein #2"
                                                                                                                                                                                                                                                                                                                                                                                                              Sequence 588 BP; 103 A; 180 C; 156 G; 149 T; 0 other;
                                                                                                                                                                                                                                                                                                                                                                                                                                            588
1195
0
0
0
0
                                                                                                                                                Goldfarb V;
                                                                                                                                                                                                                                                                                                                                                                                                                                                      Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                           sequence contains the alpha-helix0-1 variant.
                                                                                                                                                                                                                                                                                                                                                                                                                                             Length:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   US-09-965-594-12 (1-195) x AAA73329 (1-588)
                                                                                                                                                Shang Y,
                                                                                                                             ပ္ပ
                                                                                                                                                                                                                                                       Claim 26; Flg 12; 66pp; English
                                                                                                                            (BRIM ) BRISTOL-MYERS SQUIBB
                                                                                                                                                                                                                                                                                                                                                                                                                                           4.5e-84
1021.00
100.00%
100.00%
                                                                                                        99US-0115271
                                                                                    2000WO-US00345
                                                                                                                                                 Wittekind M, Weinheimer S,
                                                                                                                                                                     WPI; 2000-465976/40
                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Best Local Similarity:
                                                                                                                                                                              P-PSDB; AAB15220.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Percent Similarity:
                                           WO200040707-A1
                                                                                                                                                                                                                                                                                                                                                                                                                                  Alignment Scores:
Pred. No.:
                                                                                                        08-JAN-1999;
                                                                                    06-JAN-2000;
                                                                13-JUL-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      101
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                        Score:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           염
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      셤
  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 οχ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ó
```

```
The present sequence is the coding sequence for a mutated version of a fusion protein created using the Hepatitis C virus (HCV) NS3 and NS4A protease enzymes. These proteins are both essential for the replication of the virus, acting to cleave its replicative proteins from the polyprotein produced from the HCV genome. Inhibitors of the two proteins should be effective as antiviral treatments of HCV infection. This is useful as HCV can lead to chronic liver disease such as cirrhosis, liver failure and liver cancer. The present invention concerns a number of NS3 mutants and NS3-NS4A fusion proteins which can be used to identify inhibitors of this type, as well as enabling structural studies of the protease and protease:inhibitor complexes. The protein produced from this sequence contains the alpha-helix0-1 variant.
                                                                                                                                       SerProArgProlleSerTyrLeuLysGlySerSerGlyGlyProLeuLeuCysProAla 160
                                                                                                                                                                                                                                                                                       GlyHisAlaValGlyIlePheArgAlaAlaValCysThrArgGlyValAlaLysAlaVal 180
                                                                                                                                                                                                                                                                                                                                                           481 GGTCACGCTGTTGGTATCTTCCGTGCTGCTGTTTGCACCCGTGGTGTTGCTAAAGCTGTT 540
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Modified hepatitis C virus (HCV) NS3 protease comprising at least 1 substitution of a hydrophobic alpha-helix 0 amino acid to a hydrophilic amino acid, useful for screening inhibitors that may treat hepatitis C
                                     ThrArgHisAlaAspValIleProValArgArgArgGlyAspSerArgGlySerLeuLeu
                                                                                                                                                                             Hepatitis; NS3 protease; viral replication; chronic liver disease;
liver failure; liver cancer; mutant; mutein; ds.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Hepatitis C virus NS4A-NS3 fusion protease coding sequence #3.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                    AspPhelleProValGluSerLeuGluThrThrMetArgSerPro 195
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           /*tag= a
/product= "NS4A-NS3 fusion protein #3"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Goldfarb V;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Zhang Y,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Location/Qualifiers
1..594
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Claim 26; Fig 13; 66pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (BRIM ) BRISTOL-MYERS SQUIBB CO.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ВР
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Weinheimer S,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      AAA73330 standard; DNA; 594
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         06-JAN-2000; 2000WO-US00345
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     99US-0115271
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     19-DEC-2000 (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                WPI; 2000-465976/40.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Hepatitis C virus.
Synthetic.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      P-PSDB; AAB15221.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              WO200040707-A1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Wittekind M,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     08-JAN-1999;
```

```
P-PSDB; AAB15212
     WO200040707-A1
                                        08-JAN-1999;
                                                                                                                                                                                                                               Alignment Scores:
                13-JUL-2000
                                                                                                                                                                                                                                                                                                                  21
                                                                                                                                                                                                                                                                                                                                                                   61
                                                                                                                                                                                                                                                                                                                                                                                                        241
                                                                                                                                                                                                                                                                                                                                           41
                                                                                                                                                                                                                                                                                                                                                                                                                    101
                                                                                                                                                                                                                                                                                                                                                                                                                                             121
                                                                                                                                                                                                                                                                                                                                                       121
                                                                                                                                                                                                                                                                                                                                                                                           81
                                                                                                                                                                                                                                                                                                                                                                                                                                301
                                                                                                                                                                                                                                                             Query Match:
DB:
à
                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                  à
                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                          ö
                                                                                                                                                                                                                                                                                                                                                      셤
                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                              q
                                                                                                                                                                                                                                                                                                                                                                                           õ
                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                    ð
                                                                                                                                                                                                                                                                                                                                                                                                                                요
                                                                                                                                                                                                                                                                                                                                                                                                                                             ö
                                                                                                                                         180
                                                                                                                                                                 240
                                                                                                                                                                                         300
                                                                                                                                                                                                      118
                                                                                                                                                                                                                  360
                                                                                                                                                                                                                               138
                                                                                                                                                                                                                                          420
                                                                                                                                                                                                                                                       158
                                                                                                                                                                                                                                                                   480
                                                                                                                                                                                                                                                                                            540
                                                                                                            ProAlaGlyHisAlaValGlyIlePheArgAlaAlaValCySThrArgGlyValAlaLys 178
                                                                                         9
                                                                                                     38
                                                                                                                              58
                                                                                                                                                     78
                                                                                                                                                                              8
                                                                                                                            CGGCTCCGCAGGGTTCCCGTGCCTGACCCGTGCACCTGCGGTTCCTCCGACCTGTAC
                                                                                                                                                                                                                                                              TyralaGlnGlnThrargGlyGluGluGlyCysGlnGluThrSerGlnThrGlyArgAsp
                                                                                                                                                                                                      ProAlaProGlnGlySerArgSerLeuThrProCysThrCysGlySerSerAspLeuTyr
                                                                                                                                                                                                                               LeuValThrArgHisAlaAspValIleProValArgArgArgGlyAspSerArgGlySer
                                                                                                                                                                                                                                    LeuLeuSerProArgProIleSerTyrLeuLysGlySerSerGlyGlyProLeuLeuCys
                                                                              MetLysLysLysLysGlySerValValIleValGlyArgIleValLeuAsnGly-----Ala
                                                                                                                                                                                                                                                                                                                                                                                           protease; viral replication; chronic liver disease;
liver cancer; ds.
                                                                                                                                                                                                                                                                                                      Hepatitis C virus NS4A-NS3 fusion protease coding sequence #1
     other;
                                                                                                                                                                                                                                                                                                                                                                                                                                                   /product- "NS3-NS4A fusion protein"
                      Length:
Matches:
Conservative:
Mismatches:
Indels:
Gaps:
    T; 0
    G; 149
                                                                 (1-594)
                                                                                                                                                                                                                                                                                                                                                                                                                                 Location/Qualifiers
    C; 156
                                                                 US-09-965-594-12 (1-195) x AAA73330
                                                                                                                                                                                                                                                                                                                                           ВР
                     5.64e-82
998.00
98.48%
97.97%
    BP; 103 A; 186
                                                                                                                                                                                                                                                                                                                                           AAA73328 standard; DNA; 588
                                                                                                                                                                                                                                                                                                                                                                   (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                              ď
                                                                                                                                                                                                                                                                                                                                                                                                              Hepatitis C virus.
                                        Best Local Similarity:
Query Match:
DB:
                                                                                                                                                                                                                                                                                                                                                                                            Hepatitis; NS3
                                                                                                                                                                                                                                                                                                                                                                                                  liver failure;
                                   Percent Similarity
   Sequence 594
                                                                                                                                                                                                                                                                                                                                                                   19-DEC-2000
                Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                     Synthetic
                                                                                                      13
                                                                                                                                                                                                                  301
                                                                                                                                                                                                                               119
                                                                                                                                                                                                                                           361
                                                                                                                                                                                                                                                       139
                                                                                                                                                                                                                                                                               159
                                                                                                                                                                                                                                                                                            481
                                                                                                                                                                                                      66
                                                                                                                                                                                                                                                                                                                                                       AAA73328;
                        Pred. No.:
                                                                                                                                                                                                                                                                                                                                                                                                                                 Key
                                                                                                                                                                                                                                                                                                                               RESULT 3
                                                                                                                                                                                                                                                                                                                                     AAA73328
                                                                                                                                                                                                                                                                                                                                           g
                                                                             ò
                                                                                         g
                                                                                                      ò
                                                                                                                 셤
                                                                                                                            ò
                                                                                                                                         셤
                                                                                                                                                     ŏ
                                                                                                                                                               a
                                                                                                                                                                              ö
                                                                                                                                                                                         g
                                                                                                                                                                                                      õ
                                                                                                                                                                                                               요
                                                                                                                                                                                                                              δ
                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                QQ
                                                                                                                                                                                                                                                                               ô
                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                    õ
```

```
The present sequence is the coding sequence for a fusion protein created using the Hepatitis C virus (HCV) NS3 and NS4A protease enzymes. These proteins are both essential for the replication of the virus, acting to cleave its replicative proteins from the polyprotein produced from the HCV genome. Inhibitors of the two proteins should be effective as antiviral treatments of HCV infection. This is useful as HCV can lead to chronic liver disease such as cirrhosis, liver failure and liver cancer. The present invention concerns a number of NS3 mutants and NS3-NS4A fusion proteins which can be used to identify inhibitors of this type, as protease:inhibitor complexes.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ThrArgHisAlaAspValIleProValArgArgArgGlyAspSerArgGlySerLeuLeu 140
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    240
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         LysGlyProValIleGlnMetTyrThrAsnValAspLysAspLeuValGlyTrpProAla 100
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      300
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           80
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         GlnGlnThrArgGlyGluGluGlyCysGlnGluThrSerGlnThrGlyArgAspLysAsn 40
                                                                                                                                                                                                                          Modified hepatitis C virus (HCV) NS3 protease comprising at least 1 substitution of a hydrophobic alpha-helix 0 amino acid to a hydrophilic amino acid, useful for screening inhibitors that may treat hepatitis C
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           IleAsnGlyValCysTrpThrValTyrHisGlyAlaGlyThrArgThrIleAlaSerPro
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Sequence 588 BP; 97 A; 183 C; 153 G; 155 T; 0 other;
                                                                                                                        ,
,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Length:
Matches:
Conservative:
Mismatches:
                                                                                                                           Goldfarb
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Gaps:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     US-09-965-594-12 (1-195) x AAA73328 (1-588)
                                                                                                                        Zhang Y,
                                                                                                                                                                                                                                                                                                                                 Disclosure; Fig 10; 66pp; English.
                                                                                  ဗ
                                                                               (BRIM ) BRISTOL-MYERS SQUIBB
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1.59e-80
982.00
97.44%
97.44%
                                      99US-0115271.
                                                                                                                        Wittekind M, Weinheimer S,
06-JAN-2000; 2000WO-US00345
                                                                                                                                                                 WPI; 2000-465976/40.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Percent Similarity:
Best Local Similarity:
```

us-09-965-594-12.rng

1.99e-80

Pred. No.:

```
mutants and NS3-NS4A fusion proteins which can be used to identify inhibitors of this type, as well as enabling structural studies of the protease and protease:inhibitor complexes. The protein produced from this sequence contains the alpha-helix0-1 variant.
                                                                                                            The present sequence is the coding sequence for a mutated version of a fusion protein created using the Hepatitis C virus (HCV) NS3 and NS4A protease enzymes. These proteins are both essential for the replication of the virus, acting to cleave its replicative proteins from the polyprotein produced from the HCV genome. Inhibitors of the two proteins should be effective as antiviral treatments of HCV infection. This is useful as HCV can lead to chronic liver disease such as cirrhosis, liver failure and liver cancer. The present invention concerns a number of NS3
361 ACCCGTCACGCTGACGTTATCCCGGTTCGTCGTGGTGACTCCCGTGGTTCCCTGCTG 420
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Modified hepatitis C virus (HCV) NS3 protease comprising at least 1 substitution of a hydrophobic alpha-helix 0 amino acid to a hydrophilic amino acid, useful for screening inhibitors that may treat hepatitis C
                                                                                              GlyHisAlaValGlyIlePheArgAlaAlaValCysThrArgGlyValAlaLySAlaVal
                                                                                                                                                                                                                                                                                                                                                                                Hepatitis; NS3 protease; viral replication; chronic liver disease;
liver failure; liver cancer; mutant; mutein; ds.
                                                                                                                                                                                                                                                                                                                                                  Hepatitis C virus NS4A-NS3 fusion protease coding sequence
                                                                                                                                                           181 AspPhelleProValGluSerLeuGluThrThrMetArgSerPro 195
                                                                                                                                                                           /product= "NS4A-NS3 fusion protein #4"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Zhang Y, Goldfarb V;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Location/Qualifiers
1..594
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Claim 26; Fig 14; 66pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               (BRIM ) BRISTOL-MYERS SQUIBB CO.
                                                                                                                                                                                                                                                     BP.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Wittekind M, Weinheimer S,
                                                                                                                                                                                                                                                     AAA73331 standard; DNA; 594
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  06-JAN-2000; 2000WO-GS00345
                                                                                                                                                                                                                                                                                                                  (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            2000-465976/40.
                                                                                                                                                                                                                                                                                                                                                                                                                          Hepatitis C virus.
Synthetic.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              P-PSDB; AAB15222.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      WO200040707-A1
                                                                                                                                                                                                                                                                                                                  19-DEC-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    13-JUL-2000
                                                                                                                                                                                                                                                                                     AAA73331;
                                                             421
                                                                                            161
                                                                                                                           481
                                                                                                                                                                                                                      RESULT 4
AAA73331
ID AAA
                                                                                                                                                                                                                                                                      a
                                 ö
                                                             a
                                                                                            ö
```

Sequence 594 BP; 105 A; 187 C; 155 G; 147 T; 0 other;

Aliqnment Scores

```
LeuValThrArgHisAlaAspValIleProValArgArgArgGlyAspSerArgGlySer 138
                                                                                                                                                                                                                                                                                                                      139 LeuLeuSerProArgProIleSerTyrLeuLysGlySerSerGlyGlyProLeuLeuCys 158
                                                                           1 MetLysLysLysGlySerValVallleValGlyArglleValLeuAsnGly-----Ala 18
                                                                                     SerProLysGlyProValIleGlnMetTyrThrAsnValAspLysAspLeuValGlyTrp 98
                                                                                                                                                                                59 ThrCyslleAsnGlyValCysTrpThrValTyrHisGlyAlaGlyThrArgThrlleAla 78
                                                                                                                                                LysAsnGlnValGluGlyGluValGlnIleValSerThrAlaAlaGlnThrPheLeuAla
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 protease; viral replication; chronic liver disease; liver cancer; mutant; mutein; ds.
                                                                                                                                                                                                                                                                                                                                                                                         Hepatitis C virus NS4A-NS3 fusion protease coding sequence #8.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        /*tag== a
/product== NS4A-NS3 fusion protein #8"
594
190
               Conservative:
Mismatches:
Indels:
Length:
Watches:
                                                          US-09-965-594-12 (1-195) x AAA73331 (1-594)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                              BP
                                                                                                                                                                                                                                                                                                                                                                                                                                             AAA73335 standard; DNA; 594
                                                                                                                                                                                                                                                                                                                                                                                                                                                                               (first entry)
        981.00
96.95%
96.45%
96.08%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Hepatitis C virus.
Synthetic.
                Percent Similarity:
Best Local Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Hepatitis; NS3
liver failure;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               W0200040707-A1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                19-DEC-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 13-JUL-2000
                                                                                                                                                                                                                                                                                      119
                                                                                                                                               39
                                                                                                                                                                                                                  79
                                                                                                                                                                                                                                                                                                                                                                                                                                                              AAA73335;
                                Query Match:
                                                                                                                                                                                                                                                                                                                                                                                                                          RESULT 5
          Score:
                                                                                                                                                                                                                                                                                                                                                                                                                                              a
                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                      qq
                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                         a
                                                                                                                                                                                                                                                                                                                                                                                                         a
                                                                            ò
                                                                                             g
                                                                                                             ò
                                                                                                                            g
                                                                                                                                               õ
                                                                                                                                                               a
                                                                                                                                                                                ò
                                                                                                                                                                                                q
                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                 q
                                                                                                                                                                                                                                                    õ
                                                                                                                                                                                                                                                                                      õ
```

```
Percent Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                Wittekind M,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Sequence 594
                                                                                                                                                                                                                                                                                                                                                                                        08-JAN-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Alignment Scores
                                                                                                                                                                                                                                                                                                                                               13-JUL-2000
                                                                                                                                                                                                                                                       Synthetic
                                                  481
                              159
                                                                                                                                  ò
                                                 g
                                                                     ογ
                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                fusion protein created using the Hepatitis C virus (HCV) NS3 and NS4A protease enzymes. These proteins are both essential for the replication of the virus, acting to cleave its replicative proteins from the polyprotein produced from the HCV genome. Inhibitors of the two proteins should be effective as antiviral treatments of HCV infection. This is useful as HCV can lead to chronic liver disease such as cirrhosis, liver failure and liver cancer. The present invention concerns a number of NS3 mutants and NS3-NS4A fusion proteins which can be used to identify inhibitors of this type, as well as enabling structural studies of the protease and protease:inhibitor complexes. The protein produced from this
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            240
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CCGCCTCCGCAGGGTTCCCGTTCCCTGACCCCGTGCACCTGCGGTTCCTCCGACCTGTAC 360
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         LeuLeuSerProArgProlleSerTyrLeuLysGlySerSerGlyGlyProLeuLeuCys 158
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         78
                                                                                                                                                                                                                                                                                                                                                                                                                                                    18
                                                                                                                                                                                                                                                                                                                                                                                                                                                                       9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         38
                                                                                                                      Modified hepatitis C virus (HCV) NS3 protease comprising at least 1 substitution of a hydrophobic alpha-helix 0 amino acid to a hydrophilic amino acid, useful for screening inhibitors that may treat hepatitis C
                                                                                                                                                                                          sequence is the coding sequence for a mutated version of a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          TyrAlaGlnGlnThrArgClyGluGluGluGlyCysGlnGluThrSerGlnThrGlyArgAsp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           MetLysLysLysGlySerValValIleValGlyArgIleValLeuAsnGly-----Ala
                                                                                                                                                                                                                                                                                                                                                                                                                                                              LysasnGlnValGluGlyGluValGlnIleValSerThrAlaAlaGlnThrPheLeuAla
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ProAlaProGlnGlySerArgSerLeuThrProCysThrCysGlySerSerAspLeuTyr
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 LeuValThrArgHisAlaAspValIleProValArgArgArgGlyAspSerArgGlySer
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    sequence contains the alpha-helix0 wild-type sequence
                                                                                                                                                                                                                                                                                                                                                          594
1188
12
2
                                                                                                                                                                                                                                                                                                                            Sequence 594 BP; 98 A; 189 C; 153 G; 154 T; 0 other;
                                                                       Goldfarb V;
                                                                                                                                                                                                                                                                                                                                                         Length:
Matches:
Conservative:
Mismatches:
                                                                                                                                                                                                                                                                                                                                                                                                  Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                US-09-965-594-12 (1-195) x AAA73335 (1-594)
                                                                      Wittekind M, Weinheimer S, Zhang Y,
                                                                                                                                                                      Disclosure; Fig 18; 65pp; English.
                                                  (BRIM ) BRISTOL-MYERS SQUIBB CO
          2000WO-US00345.
                              99US-0115271
                                                                                                                                                                                                                                                                                                                                                         2e-78
959.00
95.94%
95.43%
                                                                                         2000-465976/40.
                                                                                                                                                                                                                                                                                                                                                                                        Best Local Similarity:
                                                                                                    P-PSDB; AAB15226
                                                                                                                                                                                                                                                                                                                                                                               Percent Similarity:
          06-JAN-2000;
                                                                                                                                                                                                                                                                                                                                                Alignment Scores:
                              08-JAN-1999;
                                                                                                                                                                                             present
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    241
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     361
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         139
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         29
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  119
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          19
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  39
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  79
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          66
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             301
                                                                                                                                                                                                                                                                                                                                                                                                  Query Match:
DB:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        à
a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ద
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         à
                                                                                                                                                                                                                                                                                                                                                                                                                                                    ð
```

```
mutants and NS3-NS4A fusion proteins which can be used to identify inhibitors of this type, as well as enabling structural studies of the protease and protease inhibitor complexes. The protein produced from this
ProAlaGlyHisAlaValGlyIlePheArgAlaAlaValCysThrArgGlyValAlaLys 178
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                fusion protein created using the Hepatitis C virus (HCV) NS3 and NS4A protease enzymes. These proteins are both essential for the replication of the virus, acting to cleave its replicative proteins from the polyprotein produced from the HCV genome. Inhibitors of the two proteins should be effective as antiviral treatments of HCV infection. This is useful as HCV can lead to chronic liver disease such as cirrhosis, liver failure and liver cancer. The present invention concerns a number of NS3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Modified hepatitis C virus (HCV) NS3 protease comprising at least 1 substitution of a hydrophobic alpha-helix 0 amino acid to a hydrophilic amino acid, useful for screening inhibitors that may treat hepatitis {\sf C}
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         The present sequence is the coding sequence for a mutated version of a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Hepatitis; NS3 protease; viral replication; chronic liver disease;
liver failure; liver cancer; mutant; mutein; ds.
                                                                                                                                                                                                                179 AlaValAspPheIleProValGluSerLeuGluThrThrMetArgSerPro 195
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Hepatitis C virus NS4A-NS3 fusion protease coding sequence #5.
                                                                                                                                                                                                                                                     Location/Qualiflers
1...594
/*teg a /product - "NS4A-NS3 fusion protein #5"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    other;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             594
187
1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Goldfarb V;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Length:
Matches:
Conservative:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 BP; 105 A; 189 C; 153 G; 147 T; 0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    protease and protease:inhibitor complexes. The sequence contains the alpha-helix0-1 variant
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Weinheimer S, Zhang Y,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Claim 26; Fig 15; 66pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ဗ
                                                                                                                                                                                                                                                                                                                                                                                                             ВР
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (BRIM ) BRISTOL-MYERS SQUIBB
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1.07e-77
951.00
95.43%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              99US-0115271
                                                                                                                                                                                                                                                                                                                                                                                                             AAA73332 standard; DNA; 594
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            06-JAN-2000; 2000WO-US00345
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     19-DEC-2000 (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        WPI; 2000-465976/40.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         P-PSDB; AAB15223.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        WO200040707-A1
```

```
180
                                                                                                                                                               240
                                                                                                                                                                                        LeuValThrArgHisAlaAspValIleProValArgArgArgGlyAspSerArgGlySer 138
                                                                                                                                                                                                                                                                               LeuLeuSerProArgProIleSerTyrLeuLysGlySerSerGlyGlyProLeuLeuCys 158
                                                                                                                                                                                                                                                                                                              ProAlaGlyHisAlaValGlyIlePheArgAlaAlaValCysThrArgGlyValAlaLys 178
                                                                                38
                                                                                                                28
                                                                                                                                                78
                                                                                                                                                                                86
                                                                9
                                                 TyrAlaglnGlnThrArgGlyGluGluGlyCysGlnGluThrSerGlnThrGlyArgAsp
                                                                                                                LysAsnGlnValGluGlyGluValGlnIleValSerThrAlaAlaGlnThrPheLeuAla
                                                                                                                                                 ThrCys1leAsnGlyValCysTrpThrValTyrHisGlyAlaGlyThrArgThr1leAla
                                                                                                                                                                                SerProLysGlyProVallleGlnMetTyrThrAsnValAspLysAspLeuValGlyTrp
                                                                                                                                                                                                                                                                                                                                                                                                                                                              Hepatitis; NS3 protease; viral replication; chronic liver disease;
liver failure; liver cancer; mutant; mutein; ds.
                                                                                                                                                                                                                                                                                                                                               AlaValAspPheIleProValGluSerLeuGluThrThrMetArgSerPro 195
                                                                                                                                                                                                                                                                                                                                                               591
                                                                                                                                                                                                                                                                                                                                                       Hepatitis C virus NS4A-NS3 fusion protease coding sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        /product= "NS4A-NS3 fusion protein #6"
721
Mismatches:
        Indels:
                  Gaps:
                                 US-09-965-594-12 (1-195) x AAA73332 (1-594)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Location/Qualifiers
1..594
/*tag= a
                                                                                                                                                                                                                                                                                                                                                                                               BP
                                                                                                                                                                                                                                                                                                                                                                                               AAA73333 standard; DNA; 594
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        06-JAN-2000; 2000WO-US00345
                                                                                                                                                                                                                                                                                                                                                                                                                                (first entry)
94.928
93.148
21
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Hepatitis C virus.
Synthetic.
Best Local Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       WO200040707-A1
                                                                                                                                                                                                                                                                                                                                                                                                                                19-DEC-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        13-JUL-2000
                                                                                                                                121
                                                                                                                                                29
                                                                                                                                                                                                                                                               361
                                                                                                                                                                                                                                                                                                                              481
                                                                                                                                                                                                                                                                                                                                               179
                                                                                19
                                                                                                                39
                                                                                                                                                                               4
                                                                                                                                                                                                                                               119
                                                                                                                                                                                                                                                                               139
                                                                                                                                                                                                                                                                                               421
                                                                                                                                                                                                                                                                                                              159
                                                                                                                                                                                                                                                                                                                                                                                                               AAA73333;
                                                                                                61
          Query Match
                                                                                                                                                                                                                                                                                                                                                                               RESULT 7
                                                                                                                                                                                                                                                                                                                                                                                                셤
                                                                                                                             셤
                                                                                                                                                              g
                                                                                                                                                                               ò
                                                                                                                                                                                              g
                                                                                                                                                                                                              ò
                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                               윱
                                                                                                                                                                                                                                                                            õ
                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                           Q
                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                              ద
                                                 ð
                                                              g
                                                                               à
                                                                                                             ò
                                                                                                                                               ç
```

```
The present sequence is the coding sequence for a mutated version of a fusion protein created using the Hepatitis C virus (HCV) NS3 and NS4A protease enzymes. These proteins are both essential for the replication of the virus, acting to cleave its replicative proteins from the polyprotein produced from the HCV genome. Inhibitors of the two proteins should be effective as antiviral treatments of HCV infection. This is useful as HCV can lead to chronic liver disease such as cirrhosis, liver failure and liver cancer. The present invention concerns a number of NS3 mutants and NS3-NS4A fusion proteins which can be used to identify inhibitors of this type, as well as enabling structural studies of the protease and protease inhibitor complexes. The protein produced from this sequence contains the alpha-helix0-7 variant.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           240
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     360
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            420
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 480
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                87
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         38
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   78
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       86
                                                                                                                   Modified hepatitis C virus (HCV) NS3 protease comprising at least 1 substitution of a hydrophobic alpha-helix 0 amino acid to a hydrophilic amino acid, useful for screening inhibitors that may treat hepatitis C
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1 ATGAAAAAAAAAAGGATCCGTTGTTATCGTCGGCCGTATCAACCTGTCCGGTGACACCGCT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SerProLysGlyProValileGlnMetLyrThrAsnValAspLysAspLeuValGlyTrp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CAGGCTCCGCAGGTTCCCGTTCCCTGACCCCGTGCACCTGCGGGTTCCTCCGACCTGTAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           TyralaGlnGlnThrArgGlyGluGluGlyCysGlnGluThrSerGlnThrGlyArgAsp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               LysAsnGlnValGluGlyGluValGlnIleValSerThrAlaAlaGlnThrPheLeuAla
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ProAlaProGlnGlySerArgSerLeuThrProCysThrCysGlySerSerAspLeuTyr
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1 MetLysLysLysGlySerValValIleValGlyArgIleValLeuAsnGly----Ala
                                                                                                                                                                                                                                                                                                                                                                                        Sequence 594 BP; 104 A; 191 C; 152 G; 147 T; 0 other;
                                                                                                                                                                                                                                                                                                                                                                                                                                Length:
Matches:
Conservative:
Mismatches:
                                                    Goldfarb
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Indels:
Gaps:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            US-09-965-594-12 (1-195) x AAA73333 (1-594)
                                                    Υ,
                                                    2hang
                          ႘
                                                                                                                                                                                     Claim 26; Fig 16; 66pp; English
                          (BRIM ) BRISTOL-MYERS SQUIBB
                                                                                                                                                                                                                                                                                                                                                                                                                           1.32e-76
939.00
94.92%
93.40%
99US-0115271
                                                    Weinheimer S,
                                                                             WPI; 2000-465976/40.
P-PSDB; AAB15224.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Best Local Similarity:
Query Match:
DB:
                                                                                                                                                                                                                                                                                                                                                                                                                                                          Percent Similarity:
                                                    Wittekind M,
                                                                                                                                                                                                                                                                                                                                                                                                                    Alignment Scores:
08-JAN-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              39
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         19
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   29
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      79
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           66
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     301
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 119
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          361
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      139
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              421
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                용
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          셤
ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      à
```

~

```
JS-09-965-594-12 (1-195) x AAA73334 (1-594)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                (SCHE ) SCHERING CORP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              25-JUN-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               25-JUN-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                04-DEC-2001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 179
                                                                                                                                                                                                                                                                                                                79
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                481
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ABA95615;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Hong Z,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  RESULT 9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ABA9561
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ద
                                                                                                                                                                                                                                                                                                                                                                                à
                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ő
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                            g
                                                                                                                                                                                  ò
                                                                                                                                                                                                               a
                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                              d
                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                          ద
                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                              g
                                                                                                               à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          mutants and NS3-NS4A fusion proteins which can be used to identify inhibitors of this type, as well as enabling structural studies of the protease and protease:inhibitor complexes. The protein produced from this sequence contains the alpha-halix0-7 variant.
ProAlaGlyHisAlaValGlyIlePheArgAlaAlaValCySThrArgGlyValAlaLys 178
                              CCGGCTGGTCACGCTGTTGGTATCTTCCGTGCTGTTTCCACCCGTGGTGTTGCTAAA 540
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        The present sequence is the coding sequence for a mutated version of a fusion protein created using the Hepatitis C virus (HCV) NS3 and NS4A protease enzymes. These proteins are both essential for the replication of the virus, acting to cleave its replicative proteins from the polyprotein produced from the HCV genome. Inhibitors of the two proteins should be effective as antiviral treatments of HCV infection. This is useful as HCV can lead to chronic liver disease such as cirrhosis, liver failure and liver cancer. The present invention concerns a number of NS3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Modified hepatitis C virus (HCV) NS3 protease comprising at least 1 substitution of a hydrophobic alpha-helix 0 amino acid to a hydrophilic amino acid, useful for screening inhibitors that may treat hepatitis C
                                                                                                                                                                                                                                                                                               Hepatitis; NS3 protease; viral replication; chronic liver disease;
liver failure; liver cancer; mutant; mutein; ds.
                                                                 179 AlaValAspPheIleProValGluSerLeuGluThrThrMetArgSerPro 195
                                                                                Hepatitis C virus NS4A-NS3 fusion protease coding sequence #7.
                                                                                                                                                                                                                                                                                                                                                                                                                                                 "NS4A-NS3 fusion protein #7"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Sequence 594 BP; 105 A; 192 C; 151 G; 146 T; 0 other;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                594
183
2
2
1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Goldfarb V;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Conservative:
Mismatches:
Indels:
Gaps:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Matches:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Length:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Zhang Y,
                                                                                                                                                                                                                                                                                                                                                                                                 Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Claim 26; Fig 17; 66pp; English.
                                                                                                                                                               BP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                (BRIM ) BRISTOL-MYERS SQUIBB
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1.08e-75
929.00
94.42%
92.89%
90.99%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 06-JAN-2000; 2000WO-US00345.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Wittekind M, Weinheimer S,
                                                                                                                                                               AAA73334 standard; DNA; 594
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               99US-0115271
                                                                                                                                                                                                                                (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                 /product-
                                                                                                                                                                                                                                                                                                                                                                                                                .. 594
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              WPI; 2000-465976/40.
P-PSDB; AAB15225.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Percent Similarity:
Best Local Similarity:
Query Match:
DB:
                                                                                                                                                                                                                                                                                                                                                 Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 WO200040707-A1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Alignment Scores:
Pred. No.:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 38-JAN-1999;
                                                                                                                                                                                                                                19-DEC-2000
                                                                                                                                                                                                                                                                                                                                                                Synthetic
                                                                                                                                                                                                 AAA73334;
                                481
```

```
178
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   158
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             480
                                                                                                                                                                                                                                              240
                                                                                                                                                                                                                                                                                                                   300
                                                                                       78
                                                                                                                                                                                                                                                                                86
                                  9
                                                                   38
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              LeuLeuSerProArgProIleSerTyrLeuLysGlySerSerGlyGlyProLeuLeuCys
                                  1 ATGAAAAAAAGGATCCGTTGTTATCGTCGGCCGTATCAACCTGTCCGGTGACACCGCT
                                                                   19 TyralaGlnGlnThrArgGlyGluGluGlyCysGlnGluThrSerGlnThrGlyArgAsp
                                                                                                                                                                                                                                                                                                   ProAlaGlyHisAlaValGlyIlePheArgAlaAlaValCysThrArgGlyValAlaLys
                                                                                                                                                                                                           SerProLysGlyProValIleGlnMetTyrThrAsnValAspLysAspLeuValGlyTrp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             421 CIGCIGICCCCGCGICCGAICTCCIACCIGAAAGGIICCICCGGIGGICCGCIGCIGIGC
                                                                                                                                        39 LysAsnGlnValGluGlyGluValGlnIleValSerThrAlaAlaGlnThrPheLeuAla
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   AlavalAspPheIleProValGluSerLeuGluThrThrMetArgSerPro 195
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Nucleic acid construct encoding chimeric Hepatitis C Virus (HCV)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     MetLysLysLysGlySerValValIleValGlyArgIleValLeuAsnGly--
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Pestivirus; Npro; protease; NS3; screening; ds
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Chimeric - Bovine viral diarrhea virus.
Chimeric - Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Chimeric BVDV/HCV NS3-wt sequence.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ABA95615 standard; DNA; 12734
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          99US-0344456
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           99US-0344456
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Lai VCH, Lau JYN;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            21-MAR-2002 (first entry)
```

(first entry)

28-MAR-2003

ABX15706;

```
pestivirus genome where the Npro protease gene is replaced with NS3 protease gene, useful for in vivo screening of compounds which inhibit HCV infection -
```

Example 2; Columns 17-28; 20pp; English.

The present invention relates to a nucleic acid construct encoding a chimeric Hepatitis C virus (HCV)-pestivirus genome. The construct comprises a pestivirus genome where a Npro pestivirus protease gene is replaced with a genome encoding a functional HCV NS3 protease. Furthermore, acch junction site recognised by the Npro protease is replaced with a junction site recognised by the HCV NS3 protease. The construct is useful for screening compounds that inhibit HCV in vivo by inhibiting HCV protease, where screening may be in cell culture or in an animal model. The present sequence is a chimeric clone of BVDV (bovine viral diarrhea virus)/HCV NS3-wt, which was used to illustrate the present invention.

Sequence 12734 BP; 4032 A; 2604 C; 3295 G; 2803 T; 0 other;

Length:
Matches:
Conservative:
Mismatches:
Indels: 5.05e-74 928.50 94.87% 92.31% Similarity: Percent Similarity: Alignment Scores: Query Match: Best Local

US-09-965-594-12 (1-195) x ABA95615 (1-12734)

LeuSerProArgProIleSerTyrLeuLysGlySerSerGlyGlyProLeuLeuCysPro 159 113 GGTAGTGTTGTTATTGTTGGTAGAATTGTTTTATCTGGTAGTGGTAGTATCACGGCGTAC 472 473 GCCCAGCAGACGAGAGGCCTCCTAGGGTGTAAGATCACCAGTCTGACTGGCCGGGACAAA 532 AlaGlyHisAlaValGlyIlePheArgAlaAlaValCysThrArgGlyValAlaLysAla 179 59 79 66 39 CTTTCGCCCCGGCCCATTTCCTACCTAAAAGGCTCCTCGGGGGGTCCGCTGTTGTGCCCC ProLysGlyProValIleGlnMetTyrThrAsnValAspLysAspLeuValGlyTrpPro 20 AlaGlnGlnThrArgGlyGluGluGlyCysGlnGluThrSerGlnThrGlyArgAspLys 40 AsnGlnValGluGlyGluValGlnIleValSerThrAlaAlaGlnThrPheLeuAlaThr CyslleAsnGlyValCysTrpThrValTyrHisGlyAlaGlyThrArgThrlleAlaSer ValAspPheIleProValGluSerLeuGluThrThrMetArgSer 194 9 80 100 713 773 140 833 893 180 953 S 120 a 셤 à a g ö 8 셤 셤 g õ g ò ð à ŝ ö 셤 ò

```
1.01e-74
918.50
94.36%
91.79%
                                         Best Local Similarity:
                                Percent Similarity:
Alignment Scores:
                                                       Query Match:
                               ABX15706 standard; DNA; 612 BP.
```

RESULT 10 ABX15706

antiviral prototoxophore associated DNA sequence, this sequence is described as a recombinant NS3/NS4 fusion protein in example 1 of the invention although it is clearly not a protein sequence

Sequence 612 BP; 120 A; 171 C; 191 G; 130 T; 0 other;

Length: Matches: Conservative: Mismatches:

Indels:

Gaps:

US-09-965-594-12 (1-195) x ABX15706 (1-612)

```
This invention relates to a novel synthetic viral prototoxophore domains apecific for a viral enzyme. This prototoxophore may be bound and modified by the viral enzyme. This prototoxophore may be bound and modified by the viral enzyme thus converting it to a toxophore. Also disclosed in the invention is a method for enhancing the anti-viral enter this method comprises contacting a cell, effected with a virus or is susceptible to infection, with a prototoxophore. The invention further comprises an assay to identify anti-viral agents, comprising contacting an infected cell with a candidate agent and comparing the ability of the agent to inhibit the growth or infectivity of the virus in the cell. The prototoxophores of the invention may be useful for reducing or inhibiting viral infectivity, by contacting a cell (e.g. lymphocyte, nerve cell, connective tissue cell, muscle cell or heptocyte) which is infected with a virus or is susceptible to infection with a virus, with an effective amount of the prototoxophore. The cells are cell lines an effective amount of the prototoxophore. The cells are cell lines or infected with a virus or is susceptible to infection with a virus, with an effective amount of the prototoxophore is also useful for ameliorating the severity of a viral infection in a subject, where the virus is selected from human immodeficiency virus (HW), herpes simplex virus (HS), rinnovirus and prototoxophore to the subject. The prototoxophores of the invention are cell surject and prototoxophore is a succeptible to be present sequence represents an entitivity and present sequence represents and partivity and prototoxophores of the present sequence represents and partivity and prototoxophore to the subject.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            has
                                                                                                                                                 Hepatitis C; ds; viral prototoxophore; anti-viral; tumour; virus; infection; antitumour; toxophore; human immunodeficiency virus; HIV infection; herpes simplex virus; HSV; rhinovirus; NS3 protease.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Novel synthetic viral prototoxophore for treating viral infections, toxin molety incorporated into substrate domain specific for viral enzyme, bound and modified by viral enzyme to get converted into
                                                                                                   Anti-viral synthetic prototoxophore associated DNA sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Shepard HM;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Example 1; Page 62; 66pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                            26-APR-2002; 2002WO-US13223.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Neuteboom STC,
                                                                                                                                                                                                                                                                                                                                                                                                                                                               27-APR-2001; 2001US-286893P.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  (NEWB-) NEWBIOTICS INC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      WPI; 2003-167102/16.
                                                                                                                                                                                                                                                                                                        WO200287500-A2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Cathers BE,
                                                                                                                                                                                                                                                                                                                                                           07-NOV-2002
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  toxophore
```

```
P-PSDB; AAP92041
                                                                                                                                                                                                                                                                                                                                                                                   Alignment Scores:
                          Houghton M,
                                                                                                                                                                                                                                                                                                                                   nepatitis.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                18
          (CHIR )
                                                                                                                                                                                                                                                                                                                                                                                                                         Query Match:
  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ద
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                d
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ö
                                                                                                                                                                                                                           100 AlabroGlnGlySerArgSerLeuThrProCysThrCysGlySerSerAspLeuTyrLeu 119
                                                                                                                                                                                                          GTCACGAGGCACGCCGATGTCATTCCCGTGCGCGGCGGGGGGTGATACCAGGGCCACCCTG 438
                                                79 GCCCAGCAGAAAGGGGCCTCCTAGGGTGCATAATCACCAGCCTAACTGGCCGGGACAAA 138
                                                                                                        ValThrArgHisAlaAspValIleProValArgArgArgGlyAspSerArgGlySerLeu 139
                                                                                                                                                                                                                                                            160 AlaGlyHisAlaValGlyIlePheArgAlaAlaValCysThrArgGlyValAlaLysAla 179
                                                                                                                                                                                                                                                                     Combined open reading frames of the hepatitis C virus (HCV) cDNA in clones 14i 11b, 7f, 7e, 8h, 33c, 40b, 37b, 35, 36, 81, 32, 33b, 25c, 14c,
                  19 GGTAGTGTGGTCATTGTGGGGTAGGATCATTTTGTCCGGTAGTGGTAGTATCACGGCGTAC 78
                                  AlaGlnGlnThrArgGlyGluGluGlyCysGlnGluThrSerGlnThrGlyArgAspLys 39
                                                                AsnGlnValGluGlyGluValGlnIleValSerThrAlaALaGlnThrPhcLeuAlaThr 59
                                                                                                CysileAsnGlyValCysTrpThrValTyrHisGlyAlaGlyThrArgThrIleAlaSer 79
                                                                                                                              ProLysGlyProValIleGlnMetTyrThrAsnValAspLysAspLeuValGlyTrpPro 99
                                                                                                                                                                                                                                                                                                   180 ValAspPheIleProValGluSerLeuGluThrThrMetArgSer 194
                                                                                                                                                                                                                                                                                                                                                                                                                               Hepatitis C virus; HCV; non-A, non-B hepatitis; NANBH
GlySerValValIleValGlyArglleValLeuAsnGly-
                                                                                                                                                                                                                                                                                                                                                                                                                                                              Location/Qualifiers 3..5300 /*tag= a
                                                                                                                                                                                                                                                                                                                                         AAN92097 standard; DNA; 5300 BP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     870S-0139886.
88US-0161072.
88US-0191263.
88US-0263584.
88US-0271450.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                88EP-0310922
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               87US-0122714
                                                                                                                                                                                                                                                                                                                                                                                  (first entry)
                                                                                                                                                                                                                                                                                                                                                                         (updated)
                                                                                                                                                                                                                                                                                                                                                                                                        clones 14; 11b, 7f, 7, 8f, 33f, 33g and 39c.
                                                                                                                                                                                                                                                                                                                                                                                                                                               Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                        25-MAR-2003
02-MAR-1990
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               18-NOV-1988;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               26-FEB-1988;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               31-MAY-1989
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               18-NOV-1987
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       30-DEC-1987
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     06-MAY-1988
26-0CT-1988
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      14-NOV-1988
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               EP318216-A
                                                                                                                                                                                                                                                                                                                                                         AAN92097;
                                                                40
                                                                                                9
                                                                                                                              80
                                                                                                                                                                             319
                                                                                                                                                                                                            379
                                                                                                                                                                                                                                                                           664
                                                                                                                                                                                                                                                                                                                         RESULT 11
                                                                                                                                                                                                                                                                                                                                 AAN92097
                 a
                                  ö
                                              a
                                                                ò
                                                                               q
                                                                                             õ
                                                                                                            셤
                                                                                                                             ò
                                                                                                                                          g
                                                                                                                                                          ö
                                                                                                                                                                            g
                                                                                                                                                                                              ç
                                                                                                                                                                                                          g
                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                        ద
                                                                                                                                                                                                                                                           õ
                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                       a
```

```
CORP.) (tag a) extending through clones 14i, 11b, 7f, 7e, 8h 33c, 40b, 37b, 35, 36, 81, 32, 33b, 25c, 14c, 8f, 33f, 33g and 39c of hepatitis CORTINE (ACT) COMBAL IN creating the composite sequence the following heterogeneities were considered. Clone 31c contains a sequence the following contains as the considered. Clone 33c contains a sequence the following contains as well as in 5 other overlapping clones, nuclectide #789 is a clone 33c, as well as in 5 other overlapping clones, nuclectide #789 is a clone 37b the corresponding nuclectide is an A. This heterogenity may have important ramifications for protein folding.

CO RECOIGE $1 in clone 37b the corresponding nuclectide is an A. This colone 8 is a T which may represent a cloning artifact clones is an A. Therefore the residue in clone 7 and in 3 other overlapping clones is an A. Therefore the residue in clone 8 is represented as a T clone 8 is a T. The 3 terminal sequence of clone 33f is coverlapping clones is a T. The 3 terminal sequence of clone 33f is coverlapping clones is a T. The 3 terminal sequence of clone 33f is coverlapping clones is a T. The 3 terminal sequence of clone 33f is clone 33f in clones is an A. The 3' terminus of clone 141 is depicted and instead the corresponding artifacts have been omitted and instead clones in an A pecause the corresponding artifacts have been of mitted and instead clones in an A performing artifacts have been of mitted and instead clones in an A performing artifacts have been of mitted and instead clones in an appear in probes to detect the presence of RCV nucleic acids in samples. The pollypeptide(s) it encodes could be used as infinite and passive immuno-assay reagents and vaccines and to generate antibodies useful in cloneit acids in manning in the corresponding detection of the corresponding performance of t
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CGCAGGGGCCGGGAGATACTGCTCGGGCCAGCCGATGGAATGGTCTCCAAGGGGTGGAGG 926
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             71
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           AlaalaGlnThrPheLeuAlaThrCysIleAsnGlyValCysTrpThrValTyrHisGly
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            32 ThrSerGlnThrGlyArgAspLysAsnGlnValGluGlyGluValGlnIleValSerThr
                                                                                                                                                                                                                                                                                                                                                                                                                                      is a double-stranded nucleotide sequence of the open reading frame
                                                                                                                                                                                                                                                                                                                                                Claim 3; Figure 26-1, 26-2, 26-3, 26-4, 26-5, 26-6; 139pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   3 LysLysGlySerValValIleValGly-----ArgileValLeuAsnGly---
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Sequence 5300 BP; 1047 A; 1606 C; 1515 G; 1130 T; 2 other;
                                                                                                                                                                                                                Purified hepatitis C virus
- and associated nucleic acids and polypeptide(s)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Length:
Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Updated on 25-MAR-2003 to correct PR field.) (Updated on 25-MAR-2003 to correct PI field.)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            US-09-965-594-12 (1-195) x AAN92097 (1-5300)
Kuo G;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        2.15e-71
894.50
89.71%
85.78%
87.61%
Choo QL,
                                                                                      WPI: 1989-159274/22
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Percent Similarity:
Best Local Similarity:
```

```
Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  25-MAR-2003
02-MAR-1990
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         887
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1047
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1107
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1167
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      132
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1287
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         152
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1347
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              172
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1407
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   192
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               AAN92103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                      927
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            52
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          92
                                                                                                                                                                                                                                                                                                                                                                                                                    18
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               112
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1467
                                                                                                                                                                                                                                                      Query Match:
DB:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             AAN92103
                                                                                                                                                                                                         Score:
                                                                                                                                                                                    Pred.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Ω×
       8
                                                                                                                                                                                                                                                                                                                                                                               셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                    a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     d
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 BXHHXB
                                                                                                                                                                                                                                                                                                                                                 ö
                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ç
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     à
                                                                                                                                             151
                                                                                                                                                                                                                                                                                           SerGlyGlyProLeuCysProAlaGlyHisAlaValGlyIlePheArgAlaAlaVal 171
                                                                                                                                                                                                                                                                                                                                                              CysThrArgGlyValAlaLysAlaValAspPheIleProValGluSerLeuGluThrThr 191
       6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               The sequence shows the composite cDNA sequence derived from the aligned hepatitis C virus (HCV) cDNA's in clones 141, 11b, 7f, 7e, 8h, 33c, 40b, 37b, 35, 36, 81, 32, 33b, 25c, 14c, 8f, 33f, 33g and 39c. The cDNA
92 AspLysAspLeuValGlyTrpProAlaProGlnGlySerArgSerLeuThrProCysThr
                                                                                                                                                                                                                    ArgGlyAspSerArgGlySerLeuLeuSerProArgProIleSerTyrLeuLysGlySer
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Hepatitis C virus gene - used for prodn. of polynucleotide probes, polypeptide(s) and antibodies for diagnosis, prevention and treatment of infection.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Hepatitis C virus; composite cDNA; probe; vaccine
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Disclosure; Fig. 26; 174pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Hepatitis C virus composite probe
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ВР
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CDNA; 5360
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       870S-0122714.
870S-0139886.
880S-0161072.
880S-0263584.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Kuo G;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       88GB-0027024
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                (CHIR ) CHIRON CORPORATION.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (updated)
(first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1467 ATGAGGTCCCCG 1478
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Q
                                                                                                                                                                                                                                                                                                                                                                                                                                   MetArgSerPro 195
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                3..5360
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Houghton M, Choo QL,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     WPI; 1989-215054/30
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          AAN90327 standard;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Pan troglodytes
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     18-NOV-1988;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               25-MAR-2003
11-NOV-1989
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         18-NOV-1987;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         30-DEC-1987;
26-FEB-1988;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  GB2212511-A
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  26-JUL-1989
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             26-OCT-1988
                                                                                                                                                                                                                    132
                                                                                                                                                                                                                                                                                         152
                                                                                                                                                                                                                                                                                                                                                            172
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             AAN90327;
                                                                                                                                                                                                                                                                                                                                                                                                                                 192
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      RESULT 12
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       AAN903277

110

AAN9

AAC

AAN9

AAC

AAN9

AAN9
                                                                                                                                                                             g
                                                                                                                                                                                                                                                a
                                                                                                                                                                                                                                                                                                                                                                                              q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                       셤
                                                                          ð
                                                                                                      g
                                                                                                                                             οχ
                                                                                                                                                                                                                  δ
                                                                                                                                                                                                                                                                                         à
                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                   ö
```

```
1227 recegerecressacerracessereaceasses 1286
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    131
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SerGlyGlyProLeuLeuCysProAlaGlyHisAlaValGlyIlePheArgAlaAlaVal 171
                                                                                                                                                                                                                                                                                                                                                                                             CysThrArgGlyValAlaLysAlaValAspPheIleProValGluSerLeuGluThrThr 191
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               AlaAlaGlnThrPheLeuAlaThrCysIleAsnGlyValCysTrpThrValTyrHisGly 71
                                                                                                                                                                                                                                                                                                                                                                                                                                             Combined open reading frames of the hepatitis C virus (HCV) cDNAs from
encodes antigens which react with antibodies in patients with non-A non-B hepatitis (NANBH). The CDNA can be used to design probes, or to synthesise polypeptides, which are used to diagnose HCV-induced NANBH, to raise antibodies for immunoassay or treatment, or to produce vaccines. See also AAAP9158, AAN99303-26, and AAN90328-36. (Updated on 25-MAR-2003 to correct PR field.)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CysGlySerSerAspLeuTyrLeuValThrArgHisAlaAspValIleProValArgArg
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ThrSerGlnThrGlyArgAspLysAsnGlnValGluGlyGluValGlnIleValSerThr
                                                                                                                                                                                                                                                                                                                                                                   3 LysLysGlySerValValIleValGly-----ArgIleValLeuAsnGly----
                                                                                                                                    Sequence 5360 BP; 1060 A; 1622 C; 1532 G; 1145 T; 1 other;
                                                                                                                                                                                          Length:
Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                 (1-5360)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                               US-09-965-594-12 (1-195) x AAN90327
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              AAN92103 standard; DNA; 6905
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ATGAGGTCCCCG 1478
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                (first entry)
                                                                                                                                                                                                              894.50
89.71%
85.78%
87.61%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             MetArgSerPro 195
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          (updated
                                                                                                                                                                                                                                Percent Similarity:
Best Local Similarity:
```

131

151

```
Combined open reading frames of the hepatitis C virus (HCV) cDNAs from clones K9-1 through 15e.
                                                 112 CysGlySerSerAspLeuTyrLeuValThrArgHisAlaAspValIleProValArgArg
   72 AlaGlyThrArgThrIleAlaSerProLysGlyProValIleGlnMetTyrThrAsnVal
                                                                                                                                                  132 ArgGlyAspSerArgGlySerLeuLeuSerProArgProIleSerTyrLeuLysGlySer
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Hepatitis C virus; HCV; non-A, non-B hepatitis; NANBH
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 and associated nucleic acids and polypeptide(s)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Claim 3; Figure 47-1 - 47-8; 139pp; English
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Location/Qualifiers 3..7310
                                                                                                                                                                                                                                                                                                                                                                     AAN92106 standard; DNA; 7310 BP.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             87US-0122714.
87US-0139886.
88US-0161072.
88US-0191263.
88US-0263584.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Kuo G;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           88US-0271450
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         88EP-0310922
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Purified hepatitis C virus
                                                                                                                                                                                                                                                                                                             1740 ATGAGGTCCCCG 1751
                                                                                                                                                                                                                                                                                                                                                                                                                     (updated)
(first entry)
                                                                                                                                                                                                                                                                                                MetArgSerPro 195
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  /*tag=
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Houghton M, Choo QL,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  WPI; 1989-159274/22
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CHIR ) CHIRON CORP.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              P-PSDB; AAP92050
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             18-NOV-1987;
30-DEC-1987;
26-FEB-1988;
06-MAY-1988;
26-OCT-1988;
14-NOV-1988;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       18-NOV-1988;
                                                                                                                                                                                                                                                                                                                                                                                                                   25-MAR-2003
02-MAR-1990
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               31-MAY-1989
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        EP318216-A.
                                                                                                                                                                                                                                                                                               192
                                                                                                                                                                                                                                                                                                                                                                                             AAN92106;
                                                                                                                                                                                                                                                                                                                                               RESULT 14
                                                                                                                                                                                                                                                                                                                                                            AAN92106
                          g
                                                                                                  ö
                                                                                                                          a
                                                                                                                                                  ö
                                                                                                                                                                       8
                                                                                                                                                                                               ò
                                                                                                                                                                                                                       셤
                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                     qq
                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                     셤
                                                                                                                                                                                                                                                                                                                                                                        õ
                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1319
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1140 CGCAGGGGCCGGGAGATACTGCTCGGGCCAGCCGATGGAATGGTCTCCAAGGGGTGGAGG 1199
                                                                                                                                                                                                                                                                                                                                                                                                                           It is a double-stranded nucleotide sequence of the open reading frame (ORF) (tag a) extending through clones 12f to 15e of hepatitis C virus (HCV) cDNA. It can be used to make oligomeric DNA hybridisation probes to detect the presence of HCV nucleic acids in samples. The polypeptide(s) it encodes could be used as immunoassay reagents and vaccines and to generate antibodies useful in diagnosis and passive immunotherapy for HCV infection/non-A, non-B hepatitis.
(Updated on 25-MAR-2003 to correct PR field.)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  21
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 3 LysLysGlySerValValIleValGly----- 17
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 -------AlaTyrAlaGlnGlnThrArgGlyGluGluGlyCysGlnGlu 31
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 32 ThrSerGlnThrGlyArgAspLysAsnGlnValGluGlyGluValGlnIleValSerThr
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Sequence 6905 BP; 1421 A; 2082 C; 1946 G; 1456 T; 0 other;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        6905
175
8
8
10
11
                          non-B hepatitis; NANBH
                                                                                                                                                                                                                                                                                                                                                                               and associated nucleic acids and polypeptide(s)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Length:
Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                                      Claim 3; Figure 32-1 - 32-7; 139pp; English
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         0S-09-965-594-12 (1-195) x AAN92103 (1-6905)
                                                                          Location/Qualifiers
                          Hepatitis C virus; HCV; non-A,
                                                                                                                                                                                                                                                                                                          ö
                                                                                                                                                                                                         870S-0139886.
88US-0161072.
88US-0191263.
88US-0263584.
88US-0271450.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     2.97e-71
894.50
89.71%
85.78%
87.61%
                                                                                                                                                                        88EP-0310922
                                                                                                                                                                                               87US-0122714
                                                                                                                                                                                                                                                                                                           Kuo
                                                                                                                                                                                                                                                                                                                                                                     Purified hepatitis C virus
                                                                                                  ø
                                                                                      3..6905
                                                                                                  /*tag=
12f through 15e
                                                                                                                                                                                                                                                                                                         Choo QL,
                                                                                                                                                                                                                                                                                                                                 WPI; 1989-159274/22
                                                                                                                                                                                                                                                                                   CORP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Percent Similarity:
Best Local Similarity:
                                                  virus
                                                                                                                                                                                                                                                                                                                                             P-PSDB; AAP92047
                                                                                                                                                                                                                                                                                  (CHIR ) CHIRON
                                                                                                                                                                                                          30-DEC-1987;
26-FEB-1988;
                                                                                                                                                                                                                                              26-OCT-1988;
14-NOV-1988;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Alignment Scores:
Pred. No.:
                                                  Hepatitis C
                                                                                                                                                                        18-NOV-1988;
                                                                                                                                                                                                                                                                                                         Houghton M,
                                                                                                                                                                                               18-NOV-1987
30-DEC-1987
                                                                                                                                                  31-MAY-1989
                                                                                                                                                                                                                                   06-MAY-1988
                                                                                                                          EP318216-A
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 18
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Query Match:
OB:
  clones
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ..
Q
                                                                          Key
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Score:
```

윤 õ ద ò g ö

õ

(updated)

```
19-JUL-2001
             01-NOV-1989
                                                                                                                                       18-NOV-1988;
                                                                                                                                                                                                                                 Houghton M,
                                                                                                                                                                                        26-OCT-1988;
                                                                                             GB2212511-A.
                                                                                                                  26-JUL-1989
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         22
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Ouery Match:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Score:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
    ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                It is a double-stranded nucleotide sequence of the open reading frame (ORF) (tag a) extending through clones K9-1 to 15e of hepatitis C virus (HCV) cDNA. It can be used to make oligomeric DNA hybridisation probes to detect the presence of HCV nucleic acids in samples. The polypeptide(s) it encodes could be used as immunoassay reagents and vaccines and to generate antibodies useful in diagnosis and passive immunotherapy for HCV infection/non-A, non-B hepatitis. (Updated on 25-MAR-2003 to correct PR field.)
                                                                                                                                                                                                                                                  1665 cccadgedecegeadaracreereggecagecargaargarrerecaagegregage 1724
                                                                                                                                                                                                                                                                                   1725 TIGCIGGCGCCCAICACGGCGTACGCCCAGCACAAGGGGCCTCCTAGGGTGCAIAAIC 1784
                                                                                                                                                                                                                                                                                                                                                                                                                                                            1965 GACCAAGACCTIGIGGGCIGGCCCGCTCCGCAAGGIAGCCGCTCATIGACACCCTGCACT 2024
                                                                                                                                                                                                                                                                                                                           AspLysAspLeuValGlyTrpProAlaProGlnGlySerArgSerLeuThrProCysThr 111
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  131
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          132 ArgGlyAspSerArgGlySerLeuLeuSerProArgProIleSerTyrLeuLysGlySer 151
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SerGlyGlyProLeuCysProAlaGlyHisAlaValGlyIlePheArgAlaAlaVal 171
                                                                                                                                                                                                                                                                                                                  51
                                                                                                                                                                                                                                                                                                                                                         52 AlaAlaGlnThrPheLeuAlaThrCysIleAsnGlyValCysTrpThrValTyrHisGly 71
                                                                                                                                                                                                                                                                        CysGlySerSerAspLeuTyrLeuValThrArgHisAlaAspValIleProValArgArg
                                                                                                                                                                                                                                                                                                                  ThrSerGlnThrGlyArgAspLysAsnGlnValGluGlyGluValGlnIleValSerThr
                                                                                                                                                                                                                                 3 LysLysGlySerValValIleValGly------ArgIleValLeuAsnGly
                                                                                                       Sequence 7310 BP; 1491 A; 2217 C; 2058 G; 1540 T; 4 other;
                                                                                                                                     Length:
Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                              US-09-965-594-12 (1-195) x AAN92106 (1-7310)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  AAN90336 standard; DNA; 7310
                                                                                                                                     3.19e-71
894.50
89.71%
85.78%
87.61%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        2265 ATGAGGTCCCCG 2276
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      MetArgSerPro 195
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            (updated)
                                                                                                                                                         Percent Similarity:
Best Local Similarity:
                                                                                                                            Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           25-MAR-2003
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        2145
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  112
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              2085
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   152
                                                                                                                                                                                                                                                                        18
                                                                                                                                                                                                                                                                                                                 32
                                                                                                                                                                                                                                                                                                                                                                                                                                          92
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      192
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     AAN90336;
                                                                                                                                                                                Query Match:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              RESULT 15
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       AAN90336
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
  88888888888888
                                                                                                                                                                                                                                                    셤
                                                                                                                                                                                                                                                                                           윱
                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                             8
                                                                                                                                                                                                                                                                                                                                                                                                                   a
                                                                                                                                                                                                                                                                                                                                                                                                                                                             셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ð
                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                  ö
                                                                                                                                                                                                                                                                                                                                   a
                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                          ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
```

```
1725 TIGCIGCGCCCATCACGGCGTACGCCCAGCAGACAAGGGGCCTCCTAGGGTGCATAATC 1784
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1785 ACCAGCCTAACTGGCGGGACAAAACCAAGTGGAGGGTGAGGTCCAGATTGTGTCAACT 1844
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           72 AlaGlyThrArgThrIleAlaSerProLysGlyProValIleGlnMetTyrThrAsnVal 91
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     18 ......AlaTyrAlaGlnGlnThrArgGlyGluGluGlyCysGlnGlu 31
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   The sequence shows a composite hepatitis C virus (HCV) cDNA, derived by aligning clones k9-1 through 15e in 5'-3' direction. The cDNA encodes artigons which react with antibodies in patients with non-A non-B hepatitis (NAMEH). The cDNA can be used to design probes, or to synthesise polypeptides, which are used to diagnose HCV-induced NAMEH, to raise antibodies for immunoassay or treatment, or to produce vaccines. See also AAP90208, and AAN90303-35.

(N.B. This record was resubmitted to correct errors in the sequence.) (Updated on 25-MAR-2003 to correct PR field.)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Hepatitis C virus gene - used for prodn. of polynucleotide probes, polypeptide(s) and antibodies for diagnosis, prevention and treatment of infection.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           32 ThrSerGluThrGlyArgAspLysAsnGluValGluGlyGluValGlnIeValSerThr
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                AlaalaGlnThrPheLeuAlaThrCysIleAsnGlyValCysTrpThrValTyrHisGly
                                                                                                        Hepatitis C virus; cDNA; clone 15e; clone k9-1; probe; vaccine; ds.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          3 LysLysGlySerValValIleValGly-----ArgIleValLeuAsnGly-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Sequence 7310 BP; 1495 A; 2218 C; 2058 G; 1539 T; 0 other;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Length:
Matches:
Conservative:
Mismatches:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Gaps:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   US-09-965-594-12 (1-195) x AAN90336 (1-7310)
                                                    Composite hepatitis C virus (HCV) cDNA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Disclosure; fig 47; 235pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Ö
                                                                                                                                                                                                                                                                                                                                                                                    87US-0122714
87US-0139886
                                                                                                                                                                                                                                                                                                                                                                                                                                           88US-0161072
88US-0263584
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      3.19e-71
                                                                                                                                                                                                                                                                                                                                  88GB-0027024
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Kuo
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               (CHIR ) CHIRON CORPORATION
(first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                894.50
89.718
85.788
87.618
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Choo QL,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        WPI; 1989-215054/30
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Percent Similarity:
Best Local Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  P-PSDB; AAP90288
                                                                                                                                                               Pan troglodytes.
                                                                                                                                                                                                                                                                                                                                                                                    18-NOV-1987;
30-DEC-1987;
26-FEB-1988;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Alignment Scores:
```

```
Db 1905 GCCGGAACGAGCACCATCGCGTCAGGGTCCTGTCATCCAGTGTATACCAATGTA 1964
Oy 92 AspLysAspLeuValGlyTrpProAlaProGlnGlySerArgSerLeuThrProCysThr 111
```

	1965 GACCAAGACCTTGTGGGCTGGCCGGTCCGCAAGGTAGCCGCTCATTGACACCCTGCACT 2024	
	À	
_	õ	
=	ွဲ့	
_	Ĕ	
=	ပ္ပ	
Ξ	ă	
=	Š	
=	S.	
_	Ļ	
	ું	
=	ភូ	
=	ဗ	
_	ပ္ပ	
=	ኟ	
= = = =	ģ	
_	š	
=	õ	
=	ខ	
=	닫	
=	Š	
=	Ř	
_	ŭ	
=	ဗ္ဗ	
=	Ç	
=	ဗ္ဗ	
=	5 F	
	ũ	
_	H	
_	ŭ	
=	ΰ	
::	\$	
<u>:</u>	2	
=	Š	
	5	
	1965 (	
	Ä,	

g &

Db Qy 9 % 9 9 %

Search completed: August 30, 2003, 19:47:47 Job time : 190.082 secs

us-09-965-594-12.rst

GenCore version 5.1.6 Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

August 30, 2003, 19:20:43 ; Search time 1890.92 Seconds Run on:

(without alignments) 2506.388 Million cell updates/sec

1021 1 MKKKGSVVIVGRIVLNGAYA.....VAKAVDFIPVESLETTMRSP 195 US-09-965-594-12 Title: Perfect score: Sequence:

0.0 0.0 0.0 0.0 BLOSUM62 Xgapop 10.0, Xgapext Ygapop 10.0, Ygapext Fgapop 6.0, Fgapext Delop 6.0, Delext Scoring table:

22781392 seqs, 12152238056 residues Searched:

hits satisfying chosen parameters: rotal number of

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0% Maximum Match 100% Listing first 45 summaries

Command line parameters:
-MODEL-frame+\_p2n.model -DEV-xlp
-MODEL-frame+\_p2n.model -DEV-xlp
-GOTGL\_1/USPTO\_SPOOL/US09965594/runat\_29082003\_151919\_28322/app\_query.fasta\_1.2872
-DB-EST -QFMT-fastap -SUFFIX-FST -MINMATCH-0.1 -LCOPCL-0 -LCOPEXT-0
-UNITS-blts -START-1 -END--1 -MATRIX-blosum62 -TRANS-bluman40.cdi -LIST-45
-UNITS-blts -START-1 -END--1 -MATRIX-blosum62 -TRANS-bluman40.cdi -LIST-45
-OUT-MIL-SCOME-pct -THR\_MAX-100 -THR\_MIN-0 -ALGEN-15 -MODE-LOCAL
-OUT-MIL-pto -NORM-ext -HEAPSIZE-500 -MINIEN-0 -MAXIEN-200000000
-NORM-PAT-PTO -NORM-ext -HEAPSIZE-50 -MINIEN-0 -MAXIEN-200000000
-NORM-PAT-PTO -NORM-EXT -NEG\_SCORES-0 -MILT -DSPBLCCK-100 -LONGLOG
-NORM-PAT-ARGEQUERY -NEG\_SCORES-0 -MAIT -DSPBLCCK-100 -LONGLOG
-DEV\_ITMEOUT-120 -WARN\_ITMEOUT-30 -THRRADS-1 -XGAPEXT-0 -XGAPEXT-0.5 -FGAPOP-6
-FGAPEXT-7 -YGAPOP-10 -YGAPEXT-0.5 -DELEXT-7

Database :

em\_estro: em\_htc: gb\_est1:\* gb\_htc:\* gb\_est3:\* gb\_est4:\* gb\_est4:\* gb\_est4:\* em\_est6uu:\* em\_estba:\* em\_esthum:\* em\_estin:\* em\_estmu:\* em\_estov:\* em\_estpl:\* EST: \* 

em\_gss\_hum:\* em\_gss\_inv:\* em\_gss\_pin:\* em\_gss\_vrt:\*

em\_gss\_fun:\* em\_gss\_mam:\* em\_gss\_mus:\*

em\_gss\_pro: em\_gss\_rod: em\_gss\_phg:\* em\_gss\_vrl:\* gb\_gssl:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

scription		BF304699 601888252	392487 AG	148820 AGENCOU	926101 AGENCOD		981847 AGENCOU	1625 BJ00162	\$121 BJ02412	5176 BJ01617	203316	3545 Mus mus	1437 HVSM	5803 AGEN	3244 9630	2274 6018	~	59585 AGEN	306 1	20860 60245206	21189 CH261-	3824	24 AU	1657	89911	93468	39745 AGENCO	62790	87051	542842	5424 AL55	211 AGENCO	96460 60275515	216999	5614 AU1	812	51426	05862 AG	72290	87	73186	55608 6030271	48430 AGENCOTIR	22016	919246 AGENCOUR	286751 CMD45_C	
DB ID		10 BF304699	BQ89248	BU14882	3 BQ92610	12 BI198486	3 BQ88184	2 BJ00162	2 BJ02412	2 BJ01617	0 BF20331	1 AK08054	0 BF63143	2 BM91580	0 BF86324	0 BF18227	0 BF307	3 BU16958	AW785806	0 BG42085	9 CC2211	2 BM80382	AU127824	3 BQ89465	0 BE28991	4 CB9934	3 BQ70974	4 CD26279	0 BG38705	3 BU54284	AL555424	3 BQ9282	2 BI19646	4 CB21699	AU125614	3 BU54081	3 BX45142	4 CD10586	3 BQ67229	3 BUSS687	3 BQ67318	2 BI75560	2 BM5	0 BE62201	3 BQ91924	4 CB28675	
• uery atch Length		0.4	٦.	0.1	6.	80.	 æ	۲.	۲.	۲.	9.	9.	'n.	٠. 	4.	₹.	₹.	₹.	٤.	۴.	٠. ٣.	m.	9.3 641	۳.	7.	7.	7	7.	.2	7	۲.	۲.	급.	٦.	<b>-</b> :	<b>-</b> .	۲.	٦.	٦.		۲.	-:	~	٦.	٦.	0.	
ŎŢ.		10	103.	10	101.	100.	100.	σ	σ	6	98.	6	97.	o	σ	6	95.	95.	σ	9	σ	σ	94	94.	6	6	6	93.	93.	93.	93.	93.	6	92.	92.	92.	92.	92.	92.	92.	92.	92.	92.	92.	92.	σ	
Sul!	1	0	C4	m		υ Ω			ω	S	c 10	11	12	0 13	7.	c 12	16	c 17	18	c 7	20	c 51	22	23	~	c 52	a	2	CI.	c 53	36	c 31	m i	;;	ž	3,5	36	m	3	35	4	C 41	42	4	0 44	4	

ALIGNMENTS

BF304699

984 bp mRNA linear EST 21-NOV-2000
6101888252F1 NIH\_MGC\_17 Homo sapiens cDNA clone IMAGE:4122276 5',
mRNA sequence.
BF304699 BF304699.1 GI:11251586 RESULT 1 BF304699/c LOCUS VERSION KEYWORDS SOURCE ORGANISM DEFINITION ACCESSION

Homo sapiens (human)
Homo sapiens
Homo sapiens
Homo sapiens
Homo sapiens
Homo sapiens
Homomalia; Metaca, Chordata; Craniata; Vertebrata; Euteleostomi;
Hommalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 984)

REFERENCE

```
/sex="male"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               103.50
37.428
26.458
10.148
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      362 c
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Percent Similarity:
Best Local Similarity:
Query Match:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    106
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              131
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   source
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   BASE COUNT
ORIGIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ..
Q
                                                               REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     FEATURES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Score:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                / Organism "Homo sapiens"
// Mol_type="mRNA"
// Mol_type="mRNA"
// Mol_type="mRNA"
// Mol_type="mRNA"
// Clone="IN-MAGE:412276"
// Lissue_type="rhabdomyosarcoma"
// Lissue_type
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                BQ892487 1199 bp mRNA linear EST 16-AUG-2002 AGENCOURT_8417538 Lupski_sympathetic_trunk Homo sapiens cDNA clone IMAGE:6192708 5', mRNA sequence.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   98 TrpProAlaProGlnGlySerArgSerLeuThr---ProCysThrCysGlySerSerAsp 116
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    117 LeuTyrLeuValThrArgHisAlaAspValIleProValArgArgArgGlyAspSerArg 136
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 137 GlySerLeuLeuSerProArgPro--------IleSerTyrLeuLysGlySer 151
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        553 GGCGCCTCTTGTGGGGAAGACCTCGATGGTGTCCAAGCTCGCGCTGCTGTACTGGAAGT 494
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               152 SerGlyGlyProLeuLeuCysProAlaGlyHisAlaValGlyIlePheArgAlaAlaVal 171
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               -----TICCAGCGCCGGGG 455
                                                                                  Contact: Robert Strausberg, Ph.D.

Email: cgapbs-remail.nih.gov
Tissue Procurement: ATCS

CDNA Library Preparation: Ling Hong/Rubin Laboratory

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Incytte Genomics, Inc.

Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov

Plate: LLCM.DOS row: g column: 13

High quality sequence stop: 646.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  NIH-MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    984
33
24
22
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Length:
Matches:
Conservative:
Mismatches:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         493 CGGCACGCTCCGGTCAGGTGCAGC-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         US-09-965-594-12 (1-195) x BF304699 (1-984)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        BQ892487
BQ892487.1 GI:22284501
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Homo sapiens (human)
Homo sapiens
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    4.35
106.00
45.24%
39.29%
10.38%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            454 TGCGCCCGAGGA 443
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  172 CysThrArgGly 175
                                                                                                                                                                                                                                                                                                                                                                                                                    1. .984
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Percent Similarity:
Best Local Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      133
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Alignment Scores:
Pred. No.:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Query Match:
DB:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      VERSION
KEYWORDS
SOURCE
ORGANISM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              DEFINITION
AUTHORS
TITLE
JOURNAL
COMMENT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   BASE COUNT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ACCESSION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      RESULT 2
                                                                                                                                                                                                                                                                                                                                                                                         FEATURES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    BQ892487
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ö
```

```
In Unpublished

Contact: Robert Strausberg, Ph.D.

Email: capabb=rémail.nih.gov

Tissue Procurement: Dr. James R. Lupski

CDNA Library Preparation: Life Technologies, Inc.

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Agencourt Bioscience Corporation

Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov

Plate: LLAM13595 row: c column: 13

High quality sequence start: 57

High quality sequence start: 57

High quality sequence stop: 394.

Location/Qualifiers

// Corganism="Homo sapiens"

// Corganism="Homo sapiens"

// Colone="Inmage"

// Colone="Inmage"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          484 TGGGATCCATTTTTAATAAAGGGTGCTCTGTTAATCATGGCGCCCACGGCCCCGCTGATA 543
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          86 GlnMetTyrThrAgnValAspLysAspLeuValGlyTrpProAlaProGlnGlySerArg 105
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 574 ---TGIGCTGCCTGCACACCCCCATGACCGATGTTGGGCTTATGTGGAAACGGCGAG 630
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ----ArgArgGlyAspSerArgGlySerLeuLeu-- 140
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                141 -----SerProArgProlleSerTyrLeuLys-------GlySerSerGl 153
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           66 TrpThrValTyrHisGlyAlaGlyThrArgThrIleAlaSerProLysGlyProValile 85
Eukaryota, Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1199)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               117 LeuTyr-LeuValThr------ArgHisAlaAspValIleProValArg----
                                                                    NIH,MGC http://mgc.nci.nlh.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               28 others
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1199
41
17
53
44
6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Length:
Matches:
Conservative:
Mismatches:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               /tissue_type="sympathetic trunk"
/dev_stage="adult, 16 yr"
/lab_host="DH10B"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               211 t
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           US-09-965-594-12 (1-195) x BQ892487 (1-1199)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               343 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Technologies."
```

ò a ð a à d ò 셤 ò q ò ద g ö δ

à

```
374 CIGGCACCITATIGGCCAGCAACAIGICCCTGAICICAGIGAGGICCICCTIGGIGAACA 315
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    255 AGGIGCCCICGGAIGGCCTIGCGCAICAIGIGTTTTTGCCCAICAGCACCACCACCACCATC 196
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  791 CGGTTGATGATAGAATGGGGTTACTGATTGCAACAGTTGAGTACCCAATCCTGCAGACAG 732
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          731 ACCCCTGGCAACATTGCCGACACCCTGCCAGGAAGGCGAGAATGCAGAGTTTCCTCTGTG 672
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               89 ThrAsnValAspLysAspLeuValGlyTrpProAlaProGlnGlySerArgSerLeuThr 108
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     129 ValArgArgArgGlyAspSer------ArgGlySerLeuLeuSerProArgPro 144
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   494 GGAGATTTTAGTGGTGATACCTAAAGCCTGGAAAAAGGAGGTCTTCTCGGCCCGAGACCA 435
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 -----SerTyrLeuLysGlySerSerGlyGly----- 154
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             434 GIGITCIGGGCIGGCACAGIGACIICAGCAIGGGGCAAIGGCACCAGCACGGGICAGCAG 375
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     -------AlaGlyHisAlaValGlyIlePhe---ArgAlaAlaVal----- 171
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   314 CAAAGCCCACATICCCCCGGATAIGAGGCAGCAGT-IICICCAGAGCIGGGIIGIIIICC 256
                                                                                                               /close_10="Interference of the property of gerald M. Rubin (University of California, Berkeley) using ZAP-CDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MCC Library."
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        52 AlaAlaGlnThrPhe---LeuAlaThrCys-----IleAsnGlyValCysTrpThrVal 68
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    12 ArgileValLeuAsnGlyAlaTyrAlaGlnGlnThrArgGlyGluGluGluCysGlnGlu 31
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ThrSerGlnThrGlyArgAspLysAsnGlnValGluGlyGluValGlnIleValSerThr 51
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   --LeuGluThrThrMetArgSerPro 195
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ProCysThrCysGlySerSerAspLeuTyrLeuValThrArgHisAlaAspValIlePro
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               554 CCCACTITGTCTCCAGGICITGAICAGCTGCACATCACTCAGGAITTCAAIGGTGCCCCT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    671 ATATCAAGCACTITCAGGGTTIGIAGAIGCIGCCATITGICGAACACCTGCIGG-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  TyrHisGlyAlaGlyThrArgThrIleAlaSerProLysGlyProValIleGlnMetTyr
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             575 AGCAGCGTG-------GCTTCGCTGGCT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ----CysThrArgGlyValAlaLysAlaValAspPhelleProValGluSer----
                                         /clone="IMAGE:4895638"
/tissue_type="neuroblastoma"
/lab_host="DH10B (phage-resistant)"
                                                                                                                                                                                                                                                                                                                                                                                                                               844
50
35
77
11
                                                                                                                                                                                                                                                                                                                                                                                                                          Length:
Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        US-09-965-594-12 (1-195) x BI198486 (1-844)
                          /db_xref="taxon:9606"
       /mol_type-"mRNA"
                                                                                                                                                                                                                                                                                                                                                                                                                          12.1
100.50
35.56%
20.92%
9.84%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 145 Ile-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Best Local Similarity:
Query Match:
DB:
                                                                                                                                                                                                                                                                                                                                   203
                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Percent Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                       Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         155
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     160
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  69
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           109
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                188
                                                                                                                                                                                                                                                                                                                                BASE COUNT
ORIGIN
                                                                                                                                                                                                                                                                                                                                                                                                                               Pred. No.:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     δy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 d
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ô
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1204 GCGGCGTGGCGAGTATATATACCGGCGAGTGCGAAGCGCACGGGGCGTGGAACGTTGACC 1145
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                BI198486 844 bp mRNA linear EST 10-JUL-2001 602760491F1 NIH_MGC_19 Homo sapiens CDNA clone IMAGE:4895638 5',
                                                                                                                                                                                                                                      1381 AGAGGGTGTGTTCANCGGTCAGGACAGGTCGCCGCACACTCGACGCGTCGGCCAGAGACT 1322
                                                                                                                                                                                                                                                                                                                              1321 TGTCGGGGGCGCGTTGGCGCATACCCCGGGTCGGATCGAGGTCAGCGCCGCTTGTATACA 1262
                                                                                                                                                                                                                                                                                                                                                                                                                          1261 GAGGGGAAA-----CAG 1250
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1249 GGGGTA---TGGTTATCACGGGCTGGGGCAGGTACT-----TCCCCTAAAGCG 1205
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1144 AA---GAGAGGCACCTGACGCCCTCCCCTGTGGGCNTGTCGATAATAACAAATGTGCAG 1088
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1087 GGCACGGIGAIGIGGITACIACGGGGCAGACCGGCICCAACGGCGCTCTCTAACAGACGC 1028
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                82 lyProvalIleGlnMetTyrThrAsnvalAspLysAspLeuValGlyTrpProAlaProG 102
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          102 InGlySerArgSerLeuThrProCysThrCysGlySerSerAspLeuTyrLeuValThrA 122
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          82
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CDNA Library Preparation: Ling Hong/Rubin Laboratory CDNA Library Preparation: Ling Hong/Rubin Laboratory CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: Incyte Genomics, Inc. Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov Plate: LLCM1781 row: f column: 23 High quality sequence stop: 794.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ----SerArgGlySer------LeuLeuSerProArgProIleSerTyrLeuLysG
                                                                                                                                                                                        LysGlySerValValIleValGlyArglleValLeuAsnGlyAlaTyrAlaGlnGlnThr
                                                                                                                                                                                                                                                                                     ---ArgGlyGluGluGlyCysGlnGluThrSerGlnThrGlyArgAspLysAsnGlnVal
                                                                                                                                                                                                                                                                                                                                                                                                                                                                      63 GlyValCysTrpThrValTyrHisGlyAlaGlyThrArgThrIleAlaSerProLys--G
                                                                                                                                                                                                                                                                                                                                                                            43 GluGlyGluValGlnIleValSerThrAlaAlaGlnThrPheLeuAlaThrCysIleAsn
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1 (bases 1 to 844)
NIH-WGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (WGC)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       122 rgHisAlaAspValIleProValArgArgArgGlyAsp-----
    46
15
15
68
7
                        Conservative:
Mismatches:
Indels:
    Matches:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    /organism="Homo sapiens"
                                                                                                                                          US-09-965-594-12 (1-195) x BQ926101 (1-1403)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              mRNA sequence.
BI198486
BI198486.1 GI:14653507
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Homo sapiens (human)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    967 GAGAAGGCGGCGGT 954
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        150 lySerSerGlyGly 154
101.50
36.75$
27.71$
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              . 844
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Homo sapiens
                     Percent Similarity:
Best Local Similarity:
Query Match:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              135
                                                                                                                                                                                                                                                                                     24
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   source
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    RESULT 5
BI198486/c
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      DEFINITION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ORGANISM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ACCESSION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             VERSION
KEYWORDS
SOURCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        FEATURES
```

S

```
BJ001625 BF01625 WF018SA CDNA Oryzias latipes CDNA clone WF018SA025C02 5',
                                                                                501
                                                                                                                        ------TrpProAlaProGlnGlySerArgSerLeuThrProCysThrCysGlySer 114
                                                                                                                                                                468
                                                                                                                                                                                                     134
                                                                                                                                                                                                                                           423
                                                                                                                                                                                                                                                                                 135 SerArgGlySerLeuLeuSerProArgProIleSerTyrLeuLysGlySerSerGlyGly 154
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Reopterygii; Reopterygii; Recepterygii; Acanthomorpha; Acanthomorpha; Acanthomorpha; Acanthomorpha; Acanthomorpha; Acharinomorpha; Beloniformes; Adrianichthyidae; Oryzinae; Oryzias.

I Chasea: I to 615)

Kohara, Y., Shin ii. Kimura, T., Narita, T., Jindo, T. and Takeda, H. Medaka EST Project in Takeda's lab
                                                                                                                                                                                                                                                                                                                           -----AAGGGAGGGGGGGGGGA 384
39 LysAsnGlnValGluGlyGluValGlnIleValSerThrAlaAlaGlnThrPheLeuAla 58
                                          97
                                        78 AlaSerProLysGlyProValIleGlnMetTyrThrAsnValAspLysAspLeuValGly
                                                                                                                                                                                                   115 SerAspLeuTyrLeuValThrArgHisAlaAspValIleProValArgArgArgGlyAsp
                                                                                                                                                                                                                                           ------GTCTCCCGACCCTTGCGGATTGTGGGGGTTACCCCGGGTCGGGGAG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             25
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   615
42
9
50
50
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            /sex="mixture of female and male"
/tissue_type="whole embryo"
/dev_stage="seqmentation stage 20 .:
/clone_lib="MFOLSSA CDNA"
166 c 165 g 144 t
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856
Fax: 81-559-81-6855
                                                                                                                                               500 GTATAACGGTGGCCGGGGCCTCGTGGTTCTCGG-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Conservative:
Mismatches:
                                                                                533 GGGGGGGAGGAGGTGGGTTAATT-------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Contact: Tadasu Shin-i
Center For Genetic Resource Information
National Institute of Genetics
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                /organism-"Oryzias latipes"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Length:
Matches:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Oryzias latipes (Japanese medaka)
Oryzias latipes
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 /moi_type="mRNA"
/strain="Hd-rR"
/db_xref="taxon:8090"
/clone="MF01SSA025C02"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              US-09-965-594-12 (1-195) x BJ001625 (1-615)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     tshini@genes.nig.ac.jp.
Location/Qualifiers
1..615
                                                                                                                                                                                                                                                                                                                         422 ACGGAGGGCGGGATTGTATCG---
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     BJ001625.1 GI:17364516
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     111.3
99.00
33.77%
27.81%
9.70%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             =
=
=
=
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 mRNA sequence.
BJ001625
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Unpublished
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Best Local Similarity:
                                                                                                                                                                                                                                                                                                                                                                 155 Pro 155
                                                                                                                                                                                                                                                                                                                                                                                           111
383 CCG 381
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     140
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Email:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Percent Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Alignment Scores:
Pred. No.:
                                                                                                                        86
                                                                                                                                                                                                                                           467
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Query Match:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 source
                                                                                                                                                                                                                                                                                                                                                                                                                                               RESULT 7
BJ001625/c
LOCUS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   VERSION
KEYWORDS
SOURCE
ORGANISM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   BASE COUNT
ORIGIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          DEFINITION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     TITLE
JOURNAL
COMMENT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        AUTHORS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ACCESSION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   REFERENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            FEATURES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ö
                                                                                                                                                                                                                                                                                                                   qq
                                        δ
                                                                           g
                                                                                                                        ò
                                                                                                                                                       g
                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                       q
                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                             Tissue Procurement: DCTD/DTP

CDNA Library Preparation: Rubin Laboratory

CDNA Library Preparation: Rubin Laboratory

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Agencourt Bioscience Corporation

Clone distribution: MCG clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

http://ange.llnl.gov

Plate: LLCM2501 row: c column: 23

High quality sequence stop: 322.

L. 1204

/ Corganism="Homo sapiens"

/ Mol_Ltype="MRNA"

/ Mol_Ltype="MRNA"

/ Ab_Lhost="Homo sapiens"

/ Clone="IMAGE:6295246"

/ Clone="IMAGE:6295246"

/ Lissue type="mRNA"

/ Ab_Lhost="BH100 (phage-resistant)"

/ Lab_host="BH100 (phage-resistant)"

/ Clone="IOTGan: Skin; Vector: pOTB7; Site_1: XhoI; Site_2:

ECORI: CDNA made by oligo-dT priming. Directionally cloned

into EcoRL/XhoI sites using the following 5' adaptor:

GGCACGAG(G). Library constructed by Ling Hong in the

laboratory of Gerald M. Rubin (university of California,

Berkeley) using ZAP-CDNA suphthesis kit (Stratagene) and

Superscript II RT (Life Technologies). Note: this is a

NIH_MGC Library."

228 430 c 265 g 280 t
                                                             HONGULA4/ 1204 DP MRNA Linear EST 16-AUG-2002 AGENCOURT_8712410 NIH_MGC_112 Homo sapiens CDNA clone IMAGE:6295246 5', mRNA sequence.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              654
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              GTGCCAAGGCGTGTGCTCACGGGGCAGTATATGGCGCCAAACCGTGGGTGCGGAACCGCC 714
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            40
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      61 IleAsnGlyValCysTrp----ThrValTyrHisGlyAlaGlyThrArgThrIle 77
                                                                                                                                                                                                                                       Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalla; Eutheria; Primates; Catarrhini; Hominidae; Homo. I (basea 1 to 1204)
NIH-MGC http://mgc.nci.nih.gov/.
195 CCGCCAAGGGACATGCGGATCTGCTGCTTGCTTGGAGCCCACATTGTCTGCTCCC 139
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      26 ------GluGluGlyCysGlnGluThrSerGlnThrGlyArgAspLysAsn
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              41 GlnValGluGlyGluValGLnIleValSerThrAlaAlaGlnThrPheLeuAlaThrCys
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      10 ValGlyArglleValLeuAsnGlyAlaTyrAlaGlnGlnThrArgGly-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1204
17
17
58
6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Length:
Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              US-09-965-594-12 (1-195) x BQB81847 (1-1204)
                                                                                                                                                           80881847.1 GI:22273855
                                                                                                                                                                                                   Homo sapiens (human)
Homo sapiens
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   19
100.50
36.028
25.478
9.848
                                                                                                                                                                                                                                                                                                                                              Unpublished
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Percent Similarity:
Best Local Similarity:
                                                                                                                                      BQ881847
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Query Match:
                                        RESULT 6
BQ881847/c
                                                                                LOCUS
DEFINITION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ..
Q
.
                                                                                                                                                                                                                        ORGANISM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       BASE COUNT
                                                                                                                                                                                                                                                                               REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
                                                                                                                                        ACCESSION
                                                                                                                                                             VERSION
KEYWORDS
SOURCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          FEATURES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ORIGIN
```

g

ò

õ

g

셤

ò

Best Local Similarity:         27.81%         Mismatches:         50           Query Match:         9.70%         Indels:         50           DB:         12         Gaps:         7           US-09-965-594-12 (1-195) x BJ024121 (1-643)         7           QY         39 LysasnGlnValGluGlyGluValGlnIleValSerThralaAlaGlnThrPheLeuAla 58           PIIII   111   11	Oy 59 ThrCysileAsnGlyValCysTrpThrValTyTHisGlyAlaGlyThrArgThrIleAla 78	erLeuThrPro	Oy 136 ArgGlySerLeuLeuSerProArgProlleSerTyrLeuLysGlySerSer 152	RESULT 9 BJ016176 LOCUS BJ016176 LOCUS BJ016176 BJ016176 FJ01818 BJ016176 BJ016176 FJ01818 BJ016176 FJ01818 BJ016176 VERSION BJ016176 SOURCE SOURCE OVYZIAS latipes CHORDANISM C	ACLINOPERTORIL; MedDiteryOll; TeleoStel; EuteLeoStel; MedTeleoStel; ACAINTOMOPERTORIL; MedDiteryOll; TeleoStel; MedTeleoStel; Beloniformes; Adrianichthyldae; Oryzinae; Oryzias. 1 (Dases 1 to 754) AUTHORS Rohara,Y., Shin-il,Y., Khunra,T., Narita,T., Jindo,T. and Takeda,H. TILE Medaka EST Project in Takeda's lab JOURNAL Unpublished COMMENT Contact: Tadasu Shin-i Center For Genetic Resource Information National Institute of Genetics Contact: Tadasu Shin-i	1e 20
Db 511 AAAAATGACGTAGAACCAAAAGACACACCACCACCACGTTCTGGTTCTACGGGCT 452  Qy 59 ThrCysIleAsnGlyValCysTrpThrValTyrHisGlyAlaGlyThrArgThrIleAla 78	Oy 94 AspLeuValGlyTrpProAlaProGlnGlySerArgSerLeuThrPro	Qy         123	RESULT 8 BJ024121 LOCUS BJ024121 LOCUS BJ024121 BJ024121 BJ024121 BJ024121 BJ034121 BJ034121 BJ024121 ACCESSION BJ024121 BJ024121 BJ024121 BJ024121. GI:17377389	S ES J	Letter for Genetics resource intormation National Institute of Genetics 1111 Yata, Mishima, Shizuoka 411-8540, Japan Tel: 81-559-81-6856 Fax: 81-559-81-6855 Email: tshini@genes.nig.ac.jp. FEATURES Location/Qualifiers /organism-"oryzias latipes" //strain-"Hd-TR"	/db_xxef="taxon:8090" /clone="MF01SSA143D12" /clone="MF01SSA143D12" /clone="writture of female and male" /tissue_type="whole embryo" /dev_stage="segmentation stage 20 - 25" /clone_lib="MF01SSA cDNA" ORIGIN Alignment Scores: Pred: No.: Pred: No.: 99.00 Matches: 42 Percent Similarity: 33.77% Conservative: 9

a

ò

a

ò 셤

ò

g

ò

g ò 8

ò

a

ö

ð

```
// Abdustation of the control of cont
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Indh, M., Konno, H., Okazaki, Y., Muramafsu, M. and Hayashizaki, Y. Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         99 ProAlaProGlnGlySerArgSerLeuThrProCysThrCysGlySerSer-AspLeuTy 118
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   118 rLeuValThrArgHisAlaAspValIleProValArgArgArgGlyAspSerArgGlySe 138
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        158 ysProAlaGlyHisAlaValGlyIlePheArgAlaAlaValCysThrArgGlyValAlaL 178
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            477 Greergeagecrecratagegergreecereacerringsagergergerge 418
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ARUBUD45
Mus musculus 7 days neonate cerebellum cDNA, RIKEN full-length
enriched library, clone:A730082L10 product:weakly similar to zinc
finger protein (fragment) [Mus musculus], full insert sequence.
AK080545
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Eukaryota, Metazoa, Chordata, Craniata, Vertebrata, Euteleostomi,
Mammalia, Eutheria, Rodentia, Sciurognathi, Muridae, Murinae, Mus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      591 GACTCTGAGTGCAGGAGCGGACAGAGTGCGGATCGGAAGCGCGGGGATGCTGATGGGGG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       138 rLeuLeuSer-ProArgProlleSerTyrLeuLysGlySerSerGlyGlyProLeuLeuC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                642 CCGGCTCCTCCTCCTCGGGCA-----TGAGAATGTGGAGCTCAGAGATGCT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               961
32
11
32
7
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          High-efficiency full-length cDNA cloning Meth. Enzymol. 303, 19-44 (1999) 99279253
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Conservative:
Mismatches:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Matches:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Length:
                                    /organism="Homo sapiens"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               US-09-965-594-12 (1-195) x BF203316 (1-961)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Carninci, P. and Hayashizaki, Y.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         AK080545.1 GI:26348600
HTC; CAP trapper.
Mus musculus (house mouse)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   52.44%
39.02%
9.65%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Mus musculus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Percent Similarity:
Best Local Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        178 ys 178
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  AA 416
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  .0349636
                                                                                                                                                                                                                                                                                                                                                                                                                                                                        230
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               417
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Query Match:
      source
                                                                                                                                                                                                                                                                                                                                                                                                                                                                     BASE COUNT
ORIGIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           DEFINITION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               No.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ORGANISM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          REFERENCE
AUTHORS
TITLE
JOURNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       REFERENCE
AUTHORS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      RESULT 11
AK080545
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ACCESSION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         VERSION
KEYWORDS
SOURCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                MEDLINE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               PUBMED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                TITLE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Locus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Pred.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Ω
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              BF203316 961 bp mRNA linear EST 06-NOV-2000 601865914F1 NIH_MGC_17 Homo sapiens CDNA clone IMAGE:4098578 5',
                                                                                                                                                                                                                                                                                                                                                                                                                                                                     301
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    109
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          418
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                419 CIGCAGGAIGCGGAIGIGGGTCTGCT-----TIGGITCCTCGTCCTCGAICA 469
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    -----HisalaaspvalIleProvalargargargGlyaspSer 135
                                                                                                                                                                                                                                                                                                                                                                                                              28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            78
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                93
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Eukaryota; Metazoa: Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia: Eutheria: Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 961)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Contact: Robert Strausberg, Ph.D.

Email: cgapbs-rémail.nih.gov

Tissue Procurement: ATCS

CDNA Library Preparation: Ling Hong/Rubin Laboratory

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov

Plate: LLCMSG5 row: 1 column: 03

High quality sequence stop: 637.

Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                           39 LysAsnGlnValGluGlyGluValGlnIleValSerThrAlaAlaGlnThrPheLeuAla
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                94 AspLeuValGlyTrpProAlaProGlnGlySerArgSerLeuThrPro-------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            59 ThrCysIleAsnGlyValCysTrpThrValTyrHisGlyAlaGlyThrArgThrIleAla
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                79 SerProLys------GlyProValileGlnMetTyrThrAsnValAspLys
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 380 ------ACTCGTGGAGGACGAGGGCGTCACCCCGGAGCTGTAGG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      470 TCTTCCTCACCTGACCTTCCACATCCAGGTGTCGCCAGCGCTGGTCTGACGGGTGATGGG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               136 ArgGlySerLeuLeuSerProArg-----ProlleSerTyrLeuLysGlySerSer
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      530 AGAGGCCGGACAGCAGCAGTCGGGGGTGAATCTCTGCAGGACGTCTTCACGGCGGATCA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  242 AAAATGACGTAGAACCAAAAGACACAGATCCAACCACACATGTTCTGGTTCTACGGGCT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       344 GCTCTGCGGCGGCGAGGACCTCCTGGGCCAGTTGTG---------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              110 ------CysThrCysGlySerSerAspLeuTyrLeuValThrArg------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              NIH'MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
                                                                                                                                            754
90
50
50
70
                                                                                                                                                                                                  Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                153 GlyGlyProLeuLeuCysProAlaGlyHisAla 163
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       590 GGAGGACCGACTCGCTGCAGAGCCTCTGCTGCA 622
                              u
                                                                                                                                               Length:
Matches:
                           198
/clone_lib="MF01SSA CDNA"
                                                                                                                                                                                                                                                                                                                                                     US-09-965-594-12 (1-195) x BJ016176 (1-754)
                                 ь
                              181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           BF203316.1 GI:11096902
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Homo sapiens (human)
                                                                                                                                            14.6
99.00
33.77%
27.81%
9.70%
                              181 c
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           mRNA sequence.
BF203316
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Homo saptens
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Unpublished
                                                                                                                                                                                                                                       Similarity:
                              ď
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            123 -----
                           194
                                                                                                                                                                                                     Percent Similarity:
                                                                                                                   Aliqument Scores:
                                                                                                                                                                                                                                                               Query Match:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           LOCUS
DEFINITION
                           BASE COUNT
ORIGIN
                                                                                                                                                                                                                                       Best Local
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ORGANISM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              AUTHORS
TITLE
JOURNAL
COMMENT
                                                                                                                                                  Pred. No.:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            BF203316/c
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ACCESSION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    RESULT 10
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         REFERENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     VERSION
KEYWORDS
SOURCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           FEATURES
```

158 478 HTC 05-DEC-2002

TITLE

(Mus

```
/protein_id="Bag37940.1"
/db_xref="G1:26348601"
/db_xref="G1:26348601"
/translation="bsClpAaspGRSilitprodgFFikeKlsAarravgPgpsvafgvs
TrvgadaaAgFGQRRVGARRACRYSEGLCLSRKPRRQHVPPVGPHVYGLSGGRRIPPPAGE
AQAAGRAPQOVPHPPGRRHGTVVPPQGAAGILPAlAARQVPVGPVRGREGPRRAPRHS
PRPVPTALGFSFGGGGPAPPLLAPANGRSVGLAL"
                                                                                                                                                                                                                                                                                                                                                              /tissue_type="cerebellum"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="7 days neonate"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      372 GTGGTG------CCTCCTCAAGGA 389
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           104 SerArgSerLeuThrProCysThrCysGlySerSerAspLeuTyrLeuValThrArgHis 123
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            124 AlaAspValIleProValArg---ArgArgGlyAspSerArgGlySerLeuLeuSerPro 142
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      143 ArgProlle------SerTyrLeuLysGlySerSerGlyGlyProLeu 156
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          486 AAGCCGGTTCCTACAGCCTTGGGTTTCTCGTTTGGGCAGGCTGGGCCTGCTCCTCCCCTC 545
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               157 LeuCysProAla -- - GlyHisAlaValGlyIlePheArgAlaAlaValCysThrArgGly 175
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ---CGGCCACATGGAACA 371
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        176 ValAlaLysAlaVal -----AspPheIleProValGluSerLeuGluThrThrMet 192
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   603 GACTIGAGGGTGGTTGGCTCTGGACGCAGCTCCTCTCTGGCAGGTACACTT 659
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      44 GlyGluValGlnIleValSerThrAlaAlaGlnThrPheLeuAlaThrCysIleAsnGly
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        303 GGAGAAGCGCAGGCAGCTGGACGCGCGCGCAG------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            64 ValCysTrpThrValTyrHisGlyAlaGlyThrArgThrIleAlaSerProLysGlyPro
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     84 ValileGlnMetTyrThrAsnValAspLysAspLeuValGlyTrpProAlaProGlnGly
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            /note="unnamed protein product; putative weakly similar to zinc finger protein (fragment) musculus) (PIR[148722, evidence: FASTY, 50.7%ID, 757.68[Bngth, match=601)"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1141
46
11
52
50
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Conservative:
Mismatches:
                                                                                                                                                                                    /mo__type="mRNA"
/strain="C57BL/6J"
/db_xref="Fan70M_DB:A730082L10"
/db_xref="taxon:10090"
/clone="A730082L10"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         228 t
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Length:
Matches:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 336 ----CAAGIGCCCCAICCGCCIGGG----
              URL:http://genome.gsc.riken.go.jp/
URL:http://fantom.gsc.riken.go.jp/
Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Indels:
                                                                                                                                                       organism-"Mus musculus"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Gaps:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  US-09-965-594-12 (1-195) x AK080545 (1-1141)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         353 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                'note="putative"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     /note-"putative"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  31
98.00
35.85%
28.93%
9.60%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         316 c
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Best Local Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         244 a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Percent Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       polyA_signal
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        polyA_site
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Query Match:
                                                                                                                              source
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     BASE COUNT
ORIGIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Pred. No.:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CDS
                                                                                             FEATURES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   셤
                         Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P., Konno,H., Akiyama,J., Nishi,K., Kitsunai,T., Tashiro,H., Itoh,M., Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A., Yamamoto,R., Matsunoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K., Fujiwake,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura,S., Rawai,J., Okazaki,Y., Huramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y. RIKEN Integrated sequence analysis (RISA) system-1844-format Genome Res. 10 (11), 1757-1771 (2000)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y., Arakawa, T., Hara, A., Fukunishi, Y., Konno, H., Adachi, J., Fukuda, S., Arakawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamanaka, I., Saito, T., Osiobori, T., Bono, H., Kasukawa, T., Saito, R., Kadota, K., Matsuda, Y., Ashburner, M., Batalov, S., Casavant, T., Relischmann, W., Gaasterland, T., Gissi, C., King, B., Kochiwa, H., Kuehl, P., Lewis, S., Matsuo, Y., Mikaido, I., Fesole, G., Tomita, M., Wanner, L., Washio, T., Satuhli, F., Savaiki, R., Tomita, M., Wanner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H., Baldarelli, R., Barsh, G., Blake, J., Boffelli, D., Bojunga, N., Carninci, P., de Bonaldo, M.F., Brownstein, M.J., Bult, C., Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D., Hofmann, M., Hume, D. A., Kamiya, M., Lee, N. H., Lyons, P., Marchionni, L., Mashima, J., Mazzarelli, J., Mombaerts, P., Nordone, P., Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H., Sato, K., Schonbach, C., Seya, T., Shibata, Y., Storch, K., Schonbach, C., Seya, T., Shibata, Y., Kawaji, H., Kohtsuki, S., and Hayashizaki, Y.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs

L Nature 420, 563-573 (2002)

E (Dases I to 1141)

R Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanaqaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramotto, K., Hiracka, T., Hirozane, T., Hori, F., Imotani, K., Ishii, Y., Toh, M., Kaqawa, I., Kasukawa, T., Kotoh, H., Kawai, J., Kojima, Y., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakanura, M., Nomura, K., Numazaki, R., Ohno, H., Okazoki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasazume, N., Sano, H., Satoh, M., Taqami, M., Taqawa, A., Takaku-Akahira, S., Tanaka, T., Tanak
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Submitted (16-APR-2002) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Parbloration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi'ku, Yokohama, Kanagawa 230-0045, Japan (E-mail:genome-reségsc.riken.go.jp,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Division of Experimental Animal Research in Riken contributed to
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      The FANTOM Consortium and the RIKEN Genome Exploration Research
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Functional annotation of a full-length mouse cDNA collection
Nature 409 (6821), 685-690 (2001)
21085660
11217851
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           URL:http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222
Fax:81-45-503-9216)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         prepare mouse tissues.
Please visit our web site for further details.
Genome Rcs. 10 (10), 1617-1630 (2000)
20499374
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Direct Submission
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    JOURNAL
REFERENCE
AUTHORS
                                                                                                                                                                                                                                                                                                                                                                                                                                                               JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            TITLE
JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 TITLE
              JOURNAL MEDLINE
                                                                                      PUBMED
REFERENCE
AUTHORS
```

TITLE

COMMENT

g

LOCUS

```
Unpublished
                                                                            Percent Similarity:
Best Local Similarity:
  Alignment Scores:
Pred. No.:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            EST
                                                                                                                                                                                                                                                                                                                                                                                                                                                                98
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              source
                                                                                                                                 Query Match:
DB:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            RESULT 13
BM915803/c
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     DEFINITION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ORGANISM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           AUTHORS
TITLE
JOURNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ACCESSION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  VERSION
KEYWORDS
SOURCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      REFERENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    FEATURES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          COMMENT
                                                          Score:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   LOCUS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ô
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ó
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          셤
                                                                                                                                                                                                                                                             õ
                                                                                                                                                                                                                                                                                                         셤
                                                                                                                                                                                                                                                                                                                                                            á
                                                                                                                                                                                                                                                                                                                                                                                                              q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 δy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    /issue_type='Seedling shoot'
//lab_host-*Tu71212'
//lab_host-*Tu71212'
//lab_host-*Tu71212'
//lone_lib-*Hordeum vulgare seedling shoot EST library
//lone_lib-*Hordeum vulgare seedling shoot EST library
//note="Vector: lambdaZAP; Site_1: EcoR]; Site_2: Xhol;
//note="Vector: lambdaZAP; Site_1: EcoR]; Site_2: Xhol;
//seeds were surface sterilized then germinated under axenic
conditions in the dark at room temperature on filter paper
with water, nystatin and cefotaxime in covered
crystallization dishes. Five-day old seedlings were
incubated at 90% RH for 24 hr. Shoots were then harvested,
total RNA was prepared, poly(A) RNA was multified, one
primary unamplified cDNA library was made, 600000 pfu were
primary unamplified cDNA library was made, 600000 pfu were
primary unamplified cDNA library was made, 500000 pfu were
in vivo excised to give pBluescript SK(-) cDNA phagemids.
These steps were performed in the TJ Close laboratory at
the University of Callifornia, Riverside (Choi, Close,
Penton). Phagemids were plated and picked at the Clemson
University denomics Institute (CUGI) Begum, Palmer,
Frisch, Atkins and Wing). Plasmid DNA preparations, DNA
sequencing and sequence analysis were performed at CUGI
(Wingy Yu, Frisch, Henry, Simmons, Oates, Rambo, Main).
The sequence has been trimmed to remove vector sequence
above For more details on library preparation and
                                                                                                                                                                                                                                                                                                                               Trittceae; Hordeum.

1 (bases 1 to 779)
Wing, R., Close, T.J., Kleinhofs, A., Wise, R., Begum, D., Frisch, D., Yu, Y., Henry, D., Palmer, M., Rambo, T., Simmons, J., Choi, D.W., Fenton A. B.D., Oates, R. and Main, D. Bevelopment of a genetically and physically anchored EST resource for barley genomics: Morex drought-stressed seedling shoot cDNA
                                       HVSMED0015P05f Hordeum vulgare seedling shoot EST 122-OCT-2001 HVCDRA0002 (Dehydration stress) Hordeum vulgare subsp. vulgare cDNA clone HVSMED0015P05f, mRNA sequence.
                                                                                                                                                                                                                                  Hordeum vulgare subsp. vulgare
Burodeum vulgare subsp. vulgare
Eukaryota: Viridiplantae: Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta: Liliopsida: Poales; Poaceae; Pooldeae
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           http://www.genome.clemson.edu/projects/barley. To order this clone see http://www.genome.clemson.edu/orders Also see Close IJ, Wing R, Kleinhofs A, Wise R (2001)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Genetically and physically anchored EST resources for barley genomics. Barley Genetics Newsletter 31:29-30. (http://wheat.pw.usda.gov/ggpages/bgn/31/cover.html)* 216 c 311 g 105 t
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Unpublished
On Dec 19, 2000 this sequence version replaced gi:11895595.
Contact: Wing RA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              /organism="Hordeum vulgare subsp. vulgare"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           100 Jordan Hall, Clemson, SC 29634, USA
Tel: 864 656 7288
Fax: 864 656 4293
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       /db_xref-"taxon:112509"
/clone="HVSMEb0015P05f"
/tissue_type="Seedling shoot"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Clemson University Genomics Institute
Clemson University
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Seg primer: AATTAACCCTCACTAAAGGG
High quality sequence stop: 588.
Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        sequence analysis see
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             /cultivar="Morex"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   /mol_type-"mRNA"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Email: rwing@clemson.edu
Total hq bases = 455
                                                                                                                                                                                   BF631437.2 GI:13092107
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    147
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    librar
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    BASE COUNT
ORIGIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              source
                                                                                                                                                                                VERSION
KEYWORDS
SOURCE
ORGANISM
                                                                               DEFINITION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          JOURNAL
COMMENT
RESULT 12
BF631437
                                                                                                                                                                                                                                                                                                                                                              REFERENCE
AUTHORS
                                                                                                                                                           ACCESSION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 FEATURES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  TITLE
```

```
AGENCORT_6639455 NIH_MGC_41 Homo sapiens cDNA clone IMAGE:5482056 BM915803
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   289 CCGCGGCCGTTGCCGGGGCACTATACG-TCGCCGGCGCGCGCACGACGAGGAAGAAGAACGCG 347
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     115
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 116 -AspLeuTyrLeuValThrArg-----HisAlaAspVallleProVal-----ArgAr 131
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           151 rSerGly-GlyProLeuLeuCysProAla-----------------GlyHisA 163
                                                                                                                                                                                                                                                                                                                                                                                                         49 ACCCGGTGGAGGGGGGGGGCCCCGCTGCTGCTGCTGCCGGCCCGACCGGAGCCTGC 108
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                109 CCCTGTTCTGCCAGGTGGCCGCGGTGGACGCGCGCAGGGGAAGGGGGAAGCGGCTCGTGGTCG 168
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    131 gArgGlyAspSerArgGlySerLeuLeuSerProArgProlleSerTyrLeuLysGlySe 151
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              348 CTCCGGTCGGCGCTCGCGTACGACCCGGACTCCGACGCGTGGGCCGCGCTCCCGGACATG 407
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Eukaryota; Metazoa; Chordata; Cranlata; Vertebrata; Buteleostomi; Mammalia; Butheria; Primates; Catarrhini; Hominidae; Homo. [ tbase] to 1146)
NIH-WGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammallan Gene Collection (MGC)
                                                                                                                                                                                                                                                                              ğ
                                                                                                                                                                                                                                                                                                                                                                                                                                                              -------GlnMetTyrThrAsnValAspLysAspLeuValGlyTrpPro
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           GlnIleValSerThrAlaAlaGlnThrPheLeuAlaThrCysIleAsnGlyValCysTrp
                                                                                                                                                                                                                                                                                                                                                    67 ThrvalTyrHisGlyAlaGlyThrArgThrIleAlaSerProLysGlyProValIle---
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     100 AlaProGlnGlySerArgSer------LeuThrProCysThrCysGlySerSer
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CDNA Library Preparation: Rubin Laboratory
CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MG clone distribution information can
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov i column: 01
Plate: LLCMA007 row: i column: 01
High quality sequence start: 6
High quality sequence start: 6
Location/Qualifiers
779
47
14
63
30
6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        408 GCGCAGGAGCGCGACGAGCCGCGGGCTCTGCGT 442
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   163 lavalGlyIlePheArgAlaAlaValCySThrArg 174
                                                       Conservative:
Mismatches:
                              Matches:
                                                                                                          Indels:
Length:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Contact: Robert Strausberg, Ph.D.
                                                                                                                                        Gaps:
                                                                                                                                                                                        US-09-965-594-12 (1-195) x BF631437 (1-779)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Email: cgapbs-rémail.nih.gov
Tissue Procurement: DCTD/DTP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         BM915803.1 GI:19366182
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Homo sapiens (human)
21.3
97.50
39.878
30.728
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Homo sapiens
```

```
Gaps:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        US-09-965-594-12 (1-195) x BF863244 (1-701)
                                                                                                                                                                                                                                                                                                                    Email: chauser@duke.edu.
Location/Qualifiers
                                                 Chlamydomonas reinhardtii
                                                                                                                                                                                                                        Contact: Charles Hauser
DCMB Box 91000
Duke Oniversity
Durham, NC 27708-1000
Tel: 919 613 8159
Fax: 919 613 8177
                  BF863244.1 GI:12253388
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                26
96.00
40.71%
28.32%
9.40%
                                                                                                                                                                                                          Unpublished
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Best Local Similarity:
Query Match:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     173 a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Percent Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    137
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                No.:
                                                                                                                                                                                                                                                                                                                                                       source
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     COUNT
                                                                ORGANISM
   ACCESSION
                VERSION
KEYWORDS
SOURCE
                                                                                                              REFERENCE
                                                                                                                                                                                                          JOURNAL
                                                                                                                               AUTHORS
                                                                                                                                                                                                                                                                                                                                     PEATURES
                                                                                                                                                            TITLE
                                                                                                                                                                                                                           COMMENT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    DRIGIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Pred.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     BASE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        d
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    윱
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
           /mol_type="mrnm" suprements
//db_xref="taxon:9606"
/clone="inage:18482056"
/clone="inage:18482056"
/tlssue_type="amelanotic melanoma, cell line"
/lab_host="bH10B (phage-resistant)"
/clone=lib="with_MGC_41"
/clone=lib="with_MGC_41"
/note="Organ skin; Vector: pOTB7; Site_1: XhoI; Site_2:
EcoRI; cDNA made by oligo-dr priming. Directionally cloned
into EcoRI/AhoI sites using the following 5' adaptor:
GGACGAG(G). Library constructed by Ling Hong in the
laboratory of Gerald M. Rubin (University of California,
Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
Superscript II RT (Life Technologies). Note: this is a
NIH_MGC Library."
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              BF863244 701 bp mRNA linear EST 19-JAN-2001 963042C02.xl C. reinhardtii CC-1690, Stress condition I, normalized , Lambda Lap II Chlamydomonas reinhardtii cDNA, mRNA sequence.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                903
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             gHisAlaAspValileProValArgArgArgGlyAspSerArgGlySerLeuLeuSerPr 142
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       833 ...........CGGGAGGCCGGGGGGGCCCGGTAGTGTGGCGCCGGG 796
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      142 oArgProlleSerTyrLeuLysGlySerSerGlyGlyProLeuLeuCysProAlaGlyHi 162
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     sAlaValGlyIlePheArgAlaAlaValCysThrArgGlyValAlaLysAlaValAspPh 182
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     -------GGTGGGATGTTGTGG-----TATCACTTCCCGGCGCGGGGGG 958
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            103 -GlySerArgSerLeuThrProCysThrCysGlySerSerAspLeuTyrLeuValThrAr 122
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           99
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        eLeuAlaThrCysIleAsnGlyValCysTrpThrValTyrHis------GlyAl 72
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    85
                                                                                                                                                                                                                                                                                                                                                                                                                                             17 GlyAlaTyrAlaGlnGlnThrArgGlyGluGluGlyCysGlnGluThrSer-GlnThrGl 36
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        36 yargasplysasnGlnValGluGlyGluValGlnIleValSerThrAlaAlaGlnThrPh
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    aGlyThr-----LysGlyProValll
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                eGlnMetTyrThrAsnValAspLysAspLeuValGlyTrpProAlaProGln-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           867 GGGGCCAGACTTGCTTGTCGTGTTTTCTGTGGG------
                                                                                                                                                                                                                                                                                                    Length:
Matches:
Conservative:
                                                                                                                                                                                                                                                                                                                                                   Mismatches:
                                                                                                                                                                                                                                                                                                                                                                  Indels:
/organism="Homo sapiens"
                                                                                                                                                                                                                                                                                                                                                                                                               US-09-965-594-12 (1-195) x BM915803 (1-1146)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  182 e---IleProval 185
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                690 rccccrrrg 678
                                                                                                                                                                                                                                                                                                    38.9
97.00
34.05%
26.49%
9.50%
                                                                                                                                                                                                                                                                                                                                Percent Similarity:
Best Local Similarity:
                                                                                                                                                                                                                                                                                      Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           122
                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1107
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    995
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 82
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        26
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    72
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             902
                                                                                                                                                                                                                                                                                                                                                                 Query Match:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               DEFINITION
                                                                                                                                                                                                                                        BASE COUNT
ORIGIN
                                                                                                                                                                                                                                                                                                         ..
02
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  RESULT 14
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 BF863244
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                            셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       đ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 à
```

```
/db_xref="taxon:3055"
/db_xref="taxon:3055"
/dboxref="taxon:3055"
/clone_llb="C. reinhardtii CC-1690, Stress condition I,
normalized, Lambdd 2ap II"
/note="Vector: pBluescript II SK-; Site_I: EcoRI; Site_2:
XhoI; This library, constructed by John Davies and Jeffrey
McDermott, combines colvas from CC-1690 cells grown to
mid-log phase in TAP-N (30 min, lhr, 4hr). TAP-S (30 min,
lhr, 4hr), TAP-P (4hr, 12hr, 24hr), N03 to NH4 (30min,
lhr, 4hr), TAP-P (4hr, 12hr, 24hr), N03 to NH4 (30min,
lhr, 4hr) and NH4 to N03 (30min, lhr, 4hr). PolyAm mRNA was
purified from each sample, pooled and cDNA synthesized.
The cDNA was directionally cloned into lambda Zap II
(Stratagene) in the EcoRI (5) and Xhor (3') sites.
pBluescript II SK- plasmids were excised from the lambda
ZAP clones by superinfection with Exassist (Stratagene)
phage. The library was normalized using method 4 described
in Bonaldo et al (1996) Genome Research 6: 791-806."
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             125 AspVallle----- 136
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   --- GlySerLeuLeuSerProArgProIleSerTyrLeu 148
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                285 AGAAGCCCGACCGCTGGTCCTCTGGGGTCATCCGCATGCTATGCAATCTCCCGCTATCAG 344
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             85 IleGlnMetTyrThrAsnValAspLysAspLeuValGlyTrpProAlaProGlnGlySer 104
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       105 ArgSerLeuThrProCysThrCysGlySerSerAspLeuTyrLeuValThrArgHisAla 124
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      345 GAGATCATITGCATGTGGCTTTAGTCACCCCAAGAAGAGCCTGGGAGTGGGCATTTATAA 404
Chlamydomonas reinhardtii
Chlamydomonas reinhardtii
Chlamydomonas Chlorophyta; Chlorophyceae; Volvocales;
Chlamydomonadaceae; Chlamydomonas.

1 (baees 1 to 701)
Grossmap, A., Davles, J., Federspiel, N., Harris, E., Hauser, C.,
Grossmap, A., Davles, J., Federspiel, N., Silfiow, C. and Stern, D.
Analyses of the Chlamydomonas reinhardtii Genome: A Model,
Unicellular System for Analyzing Gene Function and Regulation in
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         69 TyrHisGlyAlaGlyThrArgThrIleAlaSerProLys------GlyProVal
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  231 ACAAAGTTACATACACGG-----AAGGACCAGCGCGCTTGGCCCCCTTGGAGCCG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      /organism="Chlamydomonas reinhardtii"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           /mol_type="mRNA"
/strain="CC-1690 wild type mt+ 21gr"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           701
32
14
45
22
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Length:
Matches:
Conservative:
Mismatches:
Indels:
```

```
Search completed: August 31, 2003, 04:27:24
Job time : 1896.92 secs
                                ò
                                                                                g
                                                                                                                                               ŏ
                                                                                                                                                                                           q
                                                                                                                                                                                                                                                          ô
                                                                                                                                                                                                                                                                                                                   q
                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                               BF182274 846 bp mRNA linear EST 31-OCT-2000 601804028F1 NCI_CGAP_Mam5 Mus musculus cDNA clone IMAGE:4035102 5',
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Contact: Robert Strausberg, Ph.D.

Email: cgapbs-remail.nih.gov
Tissue Procurement: Lothar Hennighausen Ph.D., Robin Humphreys
Tissue Procurement: Lothar Hennighausen Ph.D., Robin Humphreys
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The Fechnologies, Inc.
CDNA Sequencing by: Incyte Genomics, Inc.
Clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://mage.lln.igov
Plate: LLAM9308 row: g column: 07
High quality sequence stop: 696.

| Lambard Robert Robe
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   /clone_lib-"NCI_CGAP_Mam5"
/note-"Organ: mammary; Vector: pCMV-SPORT6; Site_l: Sall; Site_2: Not1; Cloned unidirectionally. Primer: Oligo dT. Library constructed by Life Technologies. Investigators providing samples: Lothar Hennighausen/Robin Humphreys, NIH"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  102
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            709 GGTTCTTCAGTCCCAAGTGGGCTGGAGAAAGTGAACACAGAGAAAGAGGACGGTCCCTCA 650
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    nGlySerArg-----SerLeuThrProCysThrCysGlySerSerAspLeuTyrLeuVa 120
405 GAAGGGGACGGAATTCCGTTTGCGGAAAAGTCAGCGCCCCCAAGGTCTGACCAAGTGCTA 464
                                                                                                                                                                                                                                                                                                                                                                           EST.

Mus musculus (house mouse)

Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 (bases 1 to 846)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  71 GlyAlaGlyThrArg-ThrIleAlaSerProLysGlyProValIleGlnMetTyrThrAs
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  NIH-MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1 others
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    846
46
13
36
31
                                                           149 LysGlySerSerGlyGlyProLeuLeuCysProAlaGly 161
                                                                                            465 CTCCAAGGCAGCATGGGAGCCTTTCGCGGTGTGCGGGT 503
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Length:
Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         210 t
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   nValAspLysAspLeuValGlyTrp-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         US-09-965-594-12 (1-195) x BF182274 (1-846)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            241 g
                                                                                                                                                                                                                                                                                                                                                    BF182274.1 GI:11060416
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      218 c
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       33
96.00
47.20%
36.80%
9.40%
                                                                                                                                                                                                                                                                                           mRNA sequence.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Percent Similarity:
Best Local Similarity:
Query Match:
DB:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Unpublished
                                                                                                                                                                                                                                                                                                                            BF182274
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         176
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   8
                                                                                                                                                                                                                                                                                                                                              VERSION
KEYWORDS
SOURCE
ORGANISM
                                                                                                                                                                             RESULT 15
BF182274/c
                                                                                                                                                                                                                                                          DEFINITION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   BASE COUNT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Pred. No.:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
                                                                                                                                                                                                                                                                                                                      ACCESSION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            FEATURES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ORIGIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ð
                                                              ò
                                                                                                                 셤
```

```
120 lThrArgHisAlaAspValIleProValArgArgArgGlyAspSerArgGlySerLeuLe 140
                                                                                                                                            509 -ACTAGACACCT--GTAATCCAGGAGGAACGCNTGGAGGAACAGGAGGACTCC---CT 556
                                                                                                            140 uSerProArgProlleSerTyrLeuLysGlySerSerGlyGlyProLeu---LeuCysPr 159
                                                                                                                                                                                                                      159 oAlaGlyHis---AlavalGlyIlePheArg------AlaAlaValCysThrAr 174
                                                                                                                                                                                                                                                                                                                                174 gGlyvalAlaLys 178
                                                                                                                                                                                                                                                                                                                                                                     438 TGGGATGACTGAG 426
```

```
Hepatitis C virus
                                                                                                                                                 August 30, 2003, 17:42:58 ; Search time 44.6227 Seconds (without alignments) 700.745 Million cell updates/sec
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1: /SIDSI/gcgdata/geneseq/geneseqp-embl/AA1980.DAT:*
2: /SIDSI/gcgdata/geneseq/geneseqp-embl/AA1981.DAT:*
3: /SIDSI/gcgdata/geneseq-embl/AA1982.DAT:*
4: /SIDSI/gcgdata/geneseq-geneseqp-embl/AA1983.DAT:*
5: /SIDSI/gcgdata/geneseq-geneseqp-embl/AA1984.DAT:*
6: /SIDSI/gcgdata/geneseq-geneseqp-embl/AA1986.DAT:*
7: /SIDSI/gcgdata/geneseq-geneseqp-embl/AA1986.DAT:*
8: /SIDSI/gcgdata/geneseq-geneseqp-embl/AA1986.DAT:*
8: /SIDSI/gcgdata/geneseq-geneseqp-embl/AA1989.DAT:*
8: /SIDSI/gcgdata/geneseq-geneseqp-embl/AA1989.DAT:*
10: /SIDSI/gcgdata/geneseqp-embl/AA1989.DAT:*
11: /SIDSI/gcgdata/geneseqp-embl/AA1999.DAT:*
12: /SIDSI/gcgdata/geneseqp-embl/AA1999.DAT:*
13: /SIDSI/gcgdata/geneseqp-embl/AA1999.DAT:*
14: /SIDSI/gcgdata/geneseqp-embl/AA1999.DAT:*
15: /SIDSI/gcgdata/geneseqp-embl/AA1999.DAT:*
16: /SIDSI/gcgdata/geneseqy-embl/AA1999.DAT:*
17: /SIDSI/gcgdata/geneseqy-embl/AA1999.DAT:*
18: /SIDSI/gcgdata/geneseqy-embl/AA1999.DAT:*
18: /SIDSI/gcgdata/geneseqy/geneseqp-embl/AA1999.DAT:*
18: /SIDSI/gcgdata/geneseqy/geneseqp-embl/AA1999.DAT:*
18: /SIDSI/gcgdata/geneseqy/geneseqp-embl/AA1999.DAT:*
18: /SIDSI/gcgdata/geneseqy/geneseqp-embl/AA1999.DAT:*
20: /SIDSI/gcgdata/geneseqy/geneseqp-embl/AA1999.DAT:*
21: /SIDSI/gcgdata/geneseqy/geneseqp-embl/AA1999.DAT:*
22: /SIDSI/gcgdata/geneseqy/geneseqp-embl/AA1999.DAT:*
23: /SIDSI/gcgdata/geneseqy/geneseqp-embl/AA2001.DAT:*
23: /SIDSI/gcgdata/geneseqy/geneseqp-embl/AA2001.DAT:*
24: /SIDSI/gcgdata/geneseqy/geneseqp-embl/AA2003.DAT:*
24: /SIDSI/gcgdata/geneseqy/geneseqp-embl/AA2003.DAT:*
                                                                                                                                                                                                                                                                                                       1 MKKKGSVVIVGRINLSGDTA.....VAKAVDFIPVESLETTMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Description
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1107863
GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Total number of hits satisfying chosen parameters:
                                                                                                                                                                                                                                                                                                                                                                                                                                        1107863 seqs, 158726573 residues
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SUMMARIES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

    protein search, using sw model

                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              AAB15222
AAB15220
AAB15226
AAB15224
AAB15224
AAB15225
AAB15225
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      AAB15221
                                                                                                                                                                                                                                                                                                                                                                                     Gapop 10.0 , Gapext 0.5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          A_Geneseq_19Jun03:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         length: 0
length: 2000000000
                                                                                                                                                                                                                                               US-09-965-594-14
1032
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 DB
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 197
195
197
197
197
197
195
                                                                                                                                                                                                                                                                                                                                                             BLOSUM62
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Ouery
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              100.0
                                                                                                                                                                                                                                                       Title:
Perfect score:
Sequence:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              sed
                                                                                                                                                                                                                                                                                                                                                                Scoring table:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Score
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Minimum DB s
Maximum DB s
                                                                                                        OM protein
                                                                                                                                                                                                                                                                                                                                                                                                                                           Searched:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Database
                                                                                                                                                         Run on:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  450000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Result
```

110 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6
---

## ALIGNMENTS

```
Modified hepatitis C virus (HCV) NS3 protease comprising at least 1
                                                                                               Hepatitis; NS3 protease; viral replication; chronic liver disease;
liver failure; liver cancer; mutant; mutein.
                                                                                                                                                                                                                                                            Goldfarb V;
                                                                            Hepatitis C virus NS4A-NS3 fusion protease #3.
                                                                                                                                                                                                                                                            Υ,
                                                                                                                                                                                                                                                           Zhang
                     Ź
                                                                                                                                                                                                                                        (BRIM ) BRISTOL-MYERS SQUIBB CO.
                    AAB15221 standard; protein; 197
                                                                                                                                                                                               06-JAN-2000; 2000WO-US00345.
                                                                                                                                                                                                                    99US-0115271
                                                                                                                                                                                                                                                          Wittekind M, Weinheimer S,
                                                          19-DEC-2000 (first entry)
                                                                                                                                                                                                                                                                              WPI; 2000-465976/40.
N-PSDB; AAA73330.
                                                                                                                           Hepatitis C virus. Synthetic.
                                                                                                                                                         WO200040707-A1
                                                                                                                                                                                                                  08-JAN-1999;
                                                                                                                                                                             13-JUL-2000
                                       AAB15221;
RESULT 1
```

~

Page

```
ö
                                                                                     The present sequence is a mutated version of a fusion protein created using the Hepatitis C virus (HCV) NS3 and NS4A protease enzymes. These proteins are both essential for the replication of the virus, acting to cleave its replicative proteins from the polyprotein produced from the HCV genome. Inhibitors of the two proteins should be effective as antiviral treatments of HCV infection. This is useful as HCV can lead to chronic liver disease such as cirrhosis, liver failure and liver cancer. The present invention concerns a number of NS3 mutants and NS3-NS4A fusion proteins which can be used to identify inhibitors of this type, as
                                                                                                                                                                                                                                                                                                                                                                                                                              TCINGVCWTVYHGAGTRTIASPRGPVIQMYTNVDKDLVGWPAPQGSRSLTPCTCGSSDLY 120
                                                                                                                                                                                                                                                                                                                                                                                                                                            LVTRHADVIPVRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVCTRGVAK 180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1 MKKKGSVVIVGRINLSGDTAYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVSTAAQTFLA 60
                                                                                                                                                                                                                                                                                                                                                                                     substitution of a hydrophobic alpha-helix {\bf 0} amino acid to a hydrophilic amino acid, useful for screening inhibitors that may treat hepatitis {\bf C}
                                                                                                                                                                                                                       well as enabling structural studies of the protease and protease:inhibitor complexes. This sequence contains the alpha-helix0-1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  protease; viral replication; chronic liver disease;
liver cancer; mutant; mutein.
                                                                                                                                                                                                                                                                                                            Length 197;
                                                                                                                                                                                                                                                                                                                                       Indels
                                                                                                                                                                                                                                                                                                          100.0%; Score 1032; DB 21; 100.0%; Pred. No. 1.1e-98;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Goldfarb V;
                                                                                                                                                                                                                                                                                                                                          ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Hepatitis C virus NS4A-NS3 fusion protease #4
                                                                                                                                                                                                                                                                                                                                       Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Zhang Y,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Ź
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 AAB15222 standard; protein; 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             8
                                                           Claim 23; Fig 13; 66pp; English.
                                                                                                                                                                                                                                                                                                                                       .;
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              181 AVDFIPVESLETTMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              AVDFIPVESLETTMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           (BRIM ) BRISTOL-MYERS SQUIBB
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 06-JAN-2000; 2000WO-US00345
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Wittekind M, Weinheimer S,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              990S-0115271
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (first entry)
                                                                                                                                                                                                                                                                                                                                       Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  WPI; 2000-465976/40.
N-PSDB; AAA73331.
                                                                                                                                                                                                                                                                                                                        al Similarity
197; Conserv
                                                                                                                                                                                                                                                                             197 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      WO200040707-A1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Hepatitis; NS3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               liver failure;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              08-JAN-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         19-DEC-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    13-JUL-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Synthetic
                                                                                                                                                                                                                                                                                Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                              61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              AAB15222;
                                                                                                                                                                                                                                                                                                          Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                        61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        181
                                                                                                                                                                                                                                                     variant
                                                                                                                                                                                                                                                                                                                           Local
                                                                                                                                                                                                                                                                                                                                      Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     RESULT 2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                AAB15222
 αg
                                                                                                                                                                                                                                                                                                                                                                                                                             ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                        a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 đ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
```

```
ö
                                                                                                    The present sequence is a mutated version of a fusion protein created using the Hepatitis C virus (HCV) NS3 and NS4A protease enzymes. These proteins are both essential for the replication of the virus, acting to cleave its replicative proteins from the polyprotein produced from the HCV genome. Inhibitors of the two proteins should be effective as antiviral treatments of they unfection. This is useful as HCV an lead to chronic liver disease such as cirrhosis, liver failure and liver cancer. The present invention concerns a number of NS3 mutants and NS3-NS4A fusion proteins which can be used to identify inhibitors of this type, as well as enabling structural studies of the protease and protease. This sequence contains the alpha-helix0-1
                                                                                                                                                                                                                                                                                                                                                                                                                                                              61 TCINGVCWTVYHGAGTRIIASPKGPVIQMYTNVDKDLVGWPAPQGSRSLTPCTCGSSDLY 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                LVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVCTRGVAK 180
                                                                                                                                                                                                                                                                                                                                                                                                                                   9
                                                                                                                                                                                                                                                                                                                                                                                                        9
Modified hepatitis C virus (HCV) NS3 protease comprising at least 1 substitution of a hydrophobic alpha-helix 0 amino acid to a hydrophilic amino acid, useful for screening inhibitors that may treat hepatitis C
                                                                                                                                                                                                                                                                                                                                                                                                      1 MKKKGSVVIVGRINLSGDTAYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVSTAAQTFLA
                                                                                                                                                                                                                                                                                                                                                                                                                      Hepatitis; NS3 protease; viral replication; chronic liver disease; liver failure; liver cancer; mutant; mutein.
                                                                                                                                                                                                                                                                                                                                                                        ö
                                                                                                                                                                                                                                                                                                                                        Length 197;
                                                                                                                                                                                                                                                                                                                                                                         Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Goldfarb V;
                                                                                                                                                                                                                                                                                                                                      Score 1015; DB 21;
Pred. No. 6.2e-97;
0; Mismatches 3;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Hepatitis C virus NS4A-NS3 fusion protease #2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Zhang Y,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             AAB15220 standard; protein; 195 AA.
                                                                            Claim 23; Fig 14; 66pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        (BRIM ) BRISTOL-MYERS SQUIBB CO.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      AVDFIPVESLETTMRSP 197
                                                                                                                                                                                                                                                                                                                                      Query Match 98.4%;
Best Local Similarity 98.5%;
Matches 194; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          06-JAN-2000; 2000WO-US00345.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Wittekind M, Weinheimer S,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          99US-0115271
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   WPI; 2000-465976/40.
N-PSDB; AAA73329.
                                                                                                                                                                                                                                                                                                           197 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             WO200040707-A1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          08-JAN-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         19-DEC-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             13-JUL-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Synthetic.
                                                                                                                                                                                                                                                                                                            Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             AAB15220;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             AAB15220
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                RESULT
 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                a
```

3

us-09-965-594-14.rag

```
The present sequence is a mutated version of a fusion protein created using the Hepatitis C virus (HCV) NS3 and NS4A protease enzymes. These proteins are both essential for the replication of the virus, acting to cleave its replicative proteins from the polyprotein produced from the HCV genome. Inhibitors of the two proteins should be effective as antiviral treatments of HCV infection. This is useful as HCV can lead to chronic liver disease such as cirrhosis, liver failure and liver cancer. The present invention concerns a number of NS3 mutants and NS3-NS4A fusion proteins which can be used to identify inhibitors of this type, as
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 121 LVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVCTRGVAK 180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      TCINGVCWTVYHGAGTRTIASPKGPVIQMYTNVDKDLVGWPAPQGSRSLTPCTCGSSDLY 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1 HKKKGSVVIVGRIVLNG--AYAQOTRGEEGCQETSQTGRDKNQVEGEVQIVSTAAQTFLA 58
             Modified hepatitis C virus (HCV) NS3 protease comprising at least 1 substitution of a hydrophobic alpha-helix 0 amino acid to a hydrophilic amino acid, useful for screening inhibitors that may treat hepatitis C
                                                                                                                                                                                                                                                                               well as enabling structural studies of the protease and protease:inhibitor complexes. This sequence contains the alpha-helix0-1
                                                                                                                                                                                                                                                                                                                                                                                                                                                1 MKKKGSVVIVGRINLSGDTAYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVSTAAQTFLA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    59 TCINGVCWTVYHGAGTRTIASPKGPVIQMYINVDKDLVGWPAPQGSRSLTPCTCGSSDLY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Hepatitis; NS3 protease; viral replication: chronic liver disease;
liver failure; liver cancer; mutant; mutein.
                                                                                                                                                                                                                                                                                                                                                                                                                 ;
                                                                                                                                                                                                                                                                                                                                                                              Score 998; DB 21; Length 195; Pred, No. 3.5e-95;
                                                                                                                                                                                                                                                                                                                                                                                                                 Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Goldfarb V;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Hepatitis C virus NS4A-NS3 fusion protease #8
                                                                                                                                                                                                                                                                                                                                                                                                                 1; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Wittekind M, Weinheimer S, Zhang Y,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Ş
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (BRIM ) BRISTOL-MYERS SQUIBB CO
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   AAB15226 standard; protein; 197
                                                                                               Claim 23; Fig 12; 66pp; English
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   AVDFIPVESLETTMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   AVDFIPVESLETTMRSP 195
                                                                                                                                                                                                                                                                                                                                                                              96.78;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        06-JAN-2000; 2000WO-US00345.
                                                                                                                                                                                                                                                                                                                                                                                                               Matches 193; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           WPI; 2000-465976/40
                                                                                                                                                                                                                                                                                                                                                                                                 Local Similarity
                                                                                                                                                                                                                                                                                                                                                 195 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    40200040707-A1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   19-DEC-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       13-JUL-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Synthetic
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  119
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   AAB15226;
                                                                                                                                                                                                                                                                                                                                                  Sequence
                                                                                                                                                                                                                                                                                                                                                                                 Query Match
                                                                                                                                                                                                                                                                                                                  variant
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
```

9

Gaps

```
180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    fusion proteins which can be used to identify inhibitors of this type, as well as enabling structural studies of the protease and protease:inhibitor complexes. This sequence contains the alpha-helix0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    120
                                                                                                                               The present sequence is a mutated version of a fusion protein created using the Hepatitis C virus (HCV) NS3 and NS4A protease enzymes. These proteins are both essential for the replication of the virus, acting to cleave its replicative proteins from the polyprotein produced from the HCV genome. Inhibitors of the two proteins should be effective as antiviral treatments of HCV infection. This is useful as HCV can lead to chronic liver disease such as cirrhosis, liver failure and liver cancer. The present invention concerns a number of NS3 mutants and NS3-NS4A
                                                                                                                                                                                                                                                                                                                                                                                                                               9
                                                                                                                                                                                                                                                                                                                                                                                                                                                          09
                         Modified hepatitis C virus (HCV) NS3 protease comprising at least 1 substitution of a hydrophobic alpha-helix \emptyset amino acid to a hydrophilic amino acid, useful for screening inhibitors that may treat hepatitis C
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         TCINGVCWTVYHGAGIRTIASPKGPVIQMYINVDKDLVGWPAPQGSRSLIPCTCGSSDLY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         LVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVCTRGVAK
                                                                                                                                                                                                                                                                                                                                                                                                                             1 MKKKGSVVIVGRINLSGDTAYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVSTAAQTFLA
                                                                                                                                                                                                                                                                                                                                                                                                                                              Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            protease; viral replication; chronic liver disease;
                                                                                                                                                                                                                                                                                                                                                                                                ö
                                                                                                                                                                                                                                                                                                                                                                  Length 197;
                                                                                                                                                                                                                                                                                                                                                                                               Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Goldfarb V;
                                                                                                                                                                                                                                                                                                                                                               96.2%; Score 993; DB 21;
97.5%; Pred. No. 1.2e-94;
tive 0; Mismatches 5;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Hepatitis C virus NS4A-NS3 fusion protease #5.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              liver failure; liver cancer; mutant; mutein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Zhang Y,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Ź
                                                                                                     Example 5; Fig 18; 66pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (BRIM ) BRISTOL-MYERS SQUIBB CO.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      AAB15223 standard; protein; 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             181 AVDEIDVESLETTMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              181 AVDFIPVESLETTMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Wittekind M. Weinheimer S.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     06-JAN-2000; 2000WO-US00345.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 99US-0115271
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (first entry)
                                                                                                                                                                                                                                                                                                                                                                               Best Local Similarity 97.5
Matches 192; Conservative
                                                                                                                                                                                                                                                                                                      wild-type sequence.
                                                                                                                                                                                                                                                                                                                                    197 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Hepatitis C virus.
N-PSDB; AAA73335
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Hepatitis; NS3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       40200040707-A1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                19-DEC-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   38-JAN-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     13-JUL-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Synthetic
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     AAB15223;
                                                                                                                                                                                                                                                                                                                                      Seguence
                                                                                                                                                                                                                                                                                                                                                                    Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           AAB15223
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           RESULT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                        a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ö
```

```
N-PSDB; AAA73333
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  WO200040707-A1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     08-JAN-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   19-DEC-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         13-JUL-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Synthetic.
                                                                                                                                                                                                                                                                                                                                                                                                                                          121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            AAB15225;
                                                                                                                                                                                                                                                                                Sequence
                                                                                                                                                                                                                                                          variant.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           AAB15225
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 RESULT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              a
   a
                                                                                                                                                                                                                                                                                                                                                                                            ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                 a
                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                .;
0
                                                                                                                The present sequence is a mutated version of a fusion protein created using the Hepailis C virus (HCV) NG3 and NS4A protease enzymes. These proteins are both essential for the replication of the virus, acting to cleave its replicative proteins from the polyprotein produced from the HCV genome. Inhibitors of the two proteins should be effective as antiviral treatments of the two proteins should be effective as antiviral treatments of the two proteins. This is useful as HCV can lead to chronic liver disease such as cirrhosis, liver failure and liver cancer. The present invention concerns a number of NG3 mutants and NG3-NS4A fusion proteins which can be used to identify inhibitors of this type, as well as enabling structural studies of the protease and protease. This sequence contains the alpha-helix0-1
                                                                                                                                                                                                                                                                                                                                                                                 61 TCINGVCWTVYHGAGTRTIASPKGPVIQMYTNVDKDLVGWPAPQGSRSLTPCTCGSSDLY 120
                                                                                                                                                                                                                                                                                                                                                                                               121 LVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVCTRGVAK 180
                                                                                                                                                                                                                                                                                                                                                                                                                                          1 MKKKGSVVIVGRINLSGDTAYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVSTAAQTFLA 60
                                                                                                                                                                                                                                                                                                                                                           1 MKKKGSVVIVGRINLSGDTAYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVSTATQTFLA 60
                                  Modified hepatitis C virus (HCV) NS3 protease comprising at least 1 substitution of a hydrophobic alpha-helix 0 amino acid to a hydrophilic amino acid, useful for screening inhibitors that may treat hepatitis {\sf C}
                                                                                                                                                                                                                                                                                                                  Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Hepatitis; NS3 protease; viral replication; chronic liver disease;
liver failure; liver cancer; mutant; mutein.
                                                                                                                                                                                                                                                                                                                .;
0
                                                                                                                                                                                                                                                                                        Score 985; DB 21; Length 197;
Pred. No. 8e-94;
0; Mismatches 6; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Weinheimer S, Zhang Y, Goldfarb V;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Hepatitis C virus NS4A-NS3 fusion protease #6.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          AAB15224 standard; protein; 197 AA
                                                                                             Claim 23; Fig 15; 66pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                           181 AVDFIPVESLETTMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       AVDFIPVESLETTMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (BRIM ) BRISTOL-MYERS SQUIBB
                                                                                                                                                                                                                                                                                        95.4%;
97.0%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   06-JAN-2000; 2000WO-US00345
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          99US-0115271
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        (first entry)
                                                                                                                                                                                                                                                                                                               191; Conservative
WPI; 2000-465976/40.
N-PSDB; AAA73332.
                                                                                                                                                                                                                                                                                                    Similarity
                                                                                                                                                                                                                                                                    197 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Hepatitis C virus.
Synthetic.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      WO200040707-A1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          08-JAN-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Wittekind M,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       19-DEC-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             13-JUL-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 AAB15224;
                                                                                                                                                                                                                                                                     Sequence
                                                                                                                                                                                                                                                                                                                                                                                                        61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  181
                                                                                                                                                                                                                                                                                         Query Match
                                                                                                                                                                                                                                              variant
                                                                                                                                                                                                                                                                                                      Local
                                                                                                                                                                                                                                                                                                                Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               AAB1522
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ठ
                                                                                                                                                                                                                                                                                                                                                           셤
                                                                                                                                                                                                                                                                                                                                                                                   ö
                                                                                                                                                                                                                                                                                                                                                                                                     셤
                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ŝ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               a
```

```
ö
                                                                                                                                                                                                                                                                                                The present sequence is a mutated version of a fusion protein created using the Hepatitis C virus (HCV) NS3 and NS4A protease enzymes. These proteins are both essential for the replication of the virus, acting to cleave its replicative proteins from the polyprotein produced from the HCV genome. Inhibitors of the two proteins should be effective as antiviral treatments of the two proteins should be effective as antiviral treatments of HCV infection. This is useful as HCV can lead to chronic liver disease such as cirrhosis, liver failure and liver cancer. The present invention concerns a number of NS3 mutants and NS3-NS4A well as enabling structural studies of the protease and protease:inhibitor complexes. This sequence contains the alpha-helix0-7
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  61 TCINGVCWTVYHGAGTRTIASPKGPVIQMYTNVDKDLVGWPAPQGSRSLTPCTCGSSDLY 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        9
                                                                                      Modified hepatitis C virus (HCV) NS3 protease comprising at least 1 substitution of a hydrophobic alpha-helix 0 amino acid to a hydrophilic amino acid, useful for screening inhibitors that may treat hepatitis C
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           LVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVCTRGVAK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1 MKKKGSVVIVGRINLSGDTAYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVSTAAQTFLA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    0; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Hepatitis; NS3 protease; viral replication; chronic liver disease; liver failure; liver cancer; mutant; mutein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Length 197;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 7;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Score 973; DB 21;
Pred. No. 1.4e-92;
2; Mismatches 7;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Hepatitis C virus NS4A-NS3 fusion protease #7.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  AAB15225 standard; protein; 197 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (BRIM ) BRISTOL-MYERS SQUIBB CO.
                                                                                                                                                                                                                                                Claim 23; Fig 16; 66pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  AVDFIPVESLETTMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       94.38;
95.48;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       99US-0115271.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          06-JAN-2000; 2000WO-US00345.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  AVDFIPVESLETTMRSP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       94.39
Best Local Similarity 95.49
Matches 188; Conservative
2000-465976/40
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           197 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Hepatitis C virus
```

S

```
Wittekind M,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             25-MAR-2003
27-JAN-1994
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              W09315193-A1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              29-JAN-1993;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     31-JAN-1992;
                                                                                                                                                                                                                                                                                                Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       AAR40120;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 vaccine.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         RESULT 9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    AAR40120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             à
 ő
                                                                                                                                                                                                                                                                                                                                                                                          a
                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                    The present sequence is a mutated version of a fusion protein created using the Hepatitis C virus (HCV) NS3 and NS4A protease enzymes. These proteins are both essential for the replication of the virus, acting to cleave its replicative proteins from the polyprotein produced from the HCV genome. Inhibitors of the two proteins should be effective as antiviral treatments of the two proteins should be effective as antiviral treatments of the two proteins; its useful as HCV can lead to chronic liver disease such as cirrhosis, liver failure and liver cancer. The present invention concerns a number of NS3 mutants and NS3-NS4A well as enabling structural studies of the protease and proteins which can be used to identify inhibitors of this type, as well as enabling structural studies of the protease and protease:inhibitor complexes. This sequence contains the alpha-helix0-7
                                                                                                                                                                                                                                                                                                                                                                                                                   TCINGVCWTVYHGAGTRTIASPKGPVIQMYTNVDKDLVGWPAPQGSRSLTPCTCGSSDLY 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                 LVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVCTRGVAK 180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                             9
                                                                                                                                                                                                                                                                                                                                                                                           9
                                                       Modified hepatitis C virus (HCV) NS3 protease comprising at least 1 substitution of a hydrophobic alpha-helix 0 amino acid to a hydrophilic amino acid, useful for screening inhibitors that may treat hepatitis C
                                                                                                                                                                                                                                                                                                                                                                                61 TSINGVLMTVYHGAGTRTIASPKGPVTQMYTNVDKDLVGWQAPQGSRSLTPCTGGSSDLY
                                                                                                                                                                                                                                                                                                                                                                    1 MKKKGSVVIVGRINLSGDTAYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVSTAAQTFLA
                                                                                                                                                                                                                                                                                                                                               Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Hepatitis; NS3 protease; viral replication; chronic liver disease; liver failure; liver cancer.
                                                                                                                                                                                                                                                                                                                                              ö
                                                                                                                                                                                                                                                                                                                    Score 963; DB 21; Length 197;
Pred. No. 1.5e-91;
2; Mismatches 8; Indels
Weinheimer S, Zhang Y, Goldfarb V;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Hepatitis C virus NS4A-NS3 fusion protease #1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               AAB15212 standard; protein; 195 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               (BRIM ) BRISTOL-MYERS SQUIBB CO.
                                                                                                                 Claim 23; Fig 17; 66pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          AVDFIPVESLETTMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               AVDFIPVESLETTMRSP 197
                                                                                                                                                                                                                                                                                                                  93.3%;
94.9%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                06-JAN-2000; 2000WO-US00345
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        99US-0115271
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             19-DEC-2000 (first entry)
                                                                                                                                                                                                                                                                                                                                            Matches 187; Conservative
                      WPI; 2000-465976/40.
N-PSDB; AAA73334.
                                                                                                                                                                                                                                                                                                                                 Sest Local Similarity
                                                                                                                                                                                                                                                                                             197 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 WO200040707-A1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        08-JAN-1999;
Wittekind M,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         13-JUL-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Synthetic
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     AAB15212;
                                                                                                                                                                                                                                                                                                Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    181
                                                                                                                                                                                                                                                                        variant.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         RESULT 8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    AAB1521
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                         ద
                                                                                                                                                                                                                                                                                                                                                                                                                 ö
                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             οχ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
```

```
61 TCINGVCWTVTHGAGTRTIASPKGPVIQMYTNVDKDLVGWPAPQGSRSLTPCTCGSSDLY 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            121 LVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVCTRGVAK 180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1 HKKRGSVVIVGRIVLNG--AYAQQTRGLLGCIITSLIGRDKNQVEGEVQIVSTAAQTFLA 58
                                                                                                                                                                                     Modified hepatitis C virus (HCV) NS3 protease comprising at least 1 substitution of a hydrophobic alpha-helix 0 amino acid to a hydrophilic amino acid, useful for screening inhibitors that may treat hepatitis C
                                                                                                                                                                                                                                                                                                                                                                                                                                              The present sequence is a fusion protein created using the Hepatitis C virus (HCV) NS3 and NS4A protease enzymes. These proteins are both essential for the replication of the virus, acting to cleave its replicative proteins from the polyprotein produced from the HCV genome. Inhibitors of the two proteins should be effective as antiviral treatments of HCV infection. This is useful as HCV can lead to chronic liver disease such as cirrhosis, liver failure and liver cancer. The present invention concerns a number of MS3 mutants and NS3-NS4A fusion proteins which can be used to identify inhibitors of this type, as well as enabling structural studies of the protease and protease:inhibitor
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                TCINGVCWTVYHGAGTRTIASPKGPVIQMYTNVDKDLVGWPAPQGSRSLTPCTCGSSDLY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1 MKKKGSVVIVGRINLSGDTAYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVSTAAQTFLA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    HCV genomic amino acid sequence isolated from infected human LG.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          5:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Hepatitis C Virus; Non-A, non-B hepatitis Virus; HCV; NANBHV; human growth hormone; HGH: secretion signal; fusion protein;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Length 195;
    Goldfarb V;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Query Match 92.9%; Score 959; DB 21; Best Local Similarity 95.4%; Pred. No. 3.9e-91; Matches 188; Conservative 1; Mismatches 6;
    Zhang Y,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           AAR40120 standard; Protein; 3011 AA.
                                                                                                                                                                                                                                                                                                                                                                              Example 2; Fig 10; 66pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         181 AVDFIPVESLETTMRSP 197
Weinheimer S,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       93WO-0S00907
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        92US-0830024
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            (updated)
(first entry)
                                                                            WPI; 2000-465976/40.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 195 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Hepatitis C Virus
                                                                                                            N-PSDB; AAA73328
```

us-09-965-594-14.rag

```
114 CGSSDLYLVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAV 173
                                                                                                                                                                                                                                                 It is the sequence encoded in the open reading frame of hepatitis C virus cDNA inserts in clones 141,m 11b, 7f, 7e, 8h, 33c, 40b, 37b, 35, 36, 31s, 32b, 25c, 14c, 8f, a33f, 33g and 39c. It is antigenic and could be used in immunoassay reagents and vaccines and to generate antibodies useful in diagnosis and passive immunotherapy for HCV infection/non-A, non-B hepatitis.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                3 KKGSVVIVGRIN------LSGDTAYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVST
                                                                                                                                                                                                                        Claim 13; Figure 26-1, 26-2, 26-3, 26-4, 26-5, 26-6; 139pp; English
                                                                                                                                                                                                                                                                                                                                                                                                                           DB 10; Length 1766;
                                                                                                                                                                                                                                                                                                                                                                                                                                                          Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Protein sequence of hepatitis c virus composite cDNA
                                                                                                                                                                                        - and associated nucleic acids and polypeptide(s)
                                                                                                                                                                                                                                                                                                                                                                                                                         87.4%; Score 901.5; DB 1
85.8%; Pred. No. 6.9e-84;
                                                                                                                                                                                                                                                                                                                                             (Updated on 25-MAR-2003 to correct PR field.) (Updated on 25-MAR-2003 to correct PI field.)
                                                                                                                                                                                                                                                                                                                                                                                                                                                        9; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              174 CTRGVAKAVDFIPVESLETTMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Ş
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       AAP90158 standard; protein; 1786
                                                                                              ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           87US-0122714.
87US-0139886.
88US-0161072.
88US-0263584.
88US-0191263.
88US-0263584.
88US-0271450.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              88GB-0027024
                                                                                               Kuo
                                                                                                                                                                          Purified hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Hepatitis C virus; vaccine
                                                                                                                                                                                                                                                                                                                                                                                                                                                        Matches 175; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (updated)
                                                                                            Houghton M, Choo QL,
                                                                                                                           WPI; 1989-159274/22.
                                                             (CHIR ) CHIRON CORP
                                                                                                                                                                                                                                                                                                                                                                                             1766 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                           Similarity
                                                                                                                                            N-PSDB; AAN92097
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Pan troglodytes
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           18-NOV-1987;
30-DEC-1987;
26-FEB-1988;
26-OCT-1988;
06-MAY-1988;
26-OCT-1988;
14-NOV-1988;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              .8-NOV-1988;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   25-MAR-2003
10-NOV-1989
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               GB2212511-A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              26-JUL-1989
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           469
                                                                                                                                                                                                                                                                                                                                                                                               Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     AAP90158;
                                                                                                                                                                                                                                                                                                                                                                                                                           Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                          Best Local
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        RESULT 11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
   ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                           ä
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1124
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              113
                                                                                                                                                                                                                                                                                                                                                                                                                                                                         53
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Sequence encoded in the hepatitis C virus (HCV) cDNA inserts in clones 14i, 11b, 7f, 7e, 8h, 33c, 40b, 37b, 35, 36, 81, 32, 33b, 25c, 14c, 8f, 33f, 33g and 39c.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                           3 KKGSVVIVGRIN-----LSGDTAYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVST
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      AAQTFLATCINGVCWTVYHGAGTRTIASPKGPVIQMYTNVDKDLVGWPAPQGSRSLTPCT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     9; Gaps
                                                                                                                                                                                                    RMA was isolated from the plasma of a HCV seropositive human (designated "LG") and cDNA was prepared from it. The cDNA was PCR amplified using specific primers with sequences based on the prototype HCV-1 CDNA sequence (GENBANK M62121). Further amplification using nested primers resulted in 7 adjacent HCV DNA fragments which could be assembled into a full-length sequence. The DNA sequence was determined and translated into the genomic amino acid sequence. Comparison of the LG genomic amino acid sequence with that from HCV-1 showed 134 amino acid differences.
                                                                                                           New plasmid pHCV-162 is a mammalian expression systems for HCVD proteins - useful for diagnosing HCV infection and as vaccines for preventing HCV infection
                                                                                                                                                                                                                                                                                                                                                                                                            DB 14; Length 3011;
                                                                                                                                                                                                                                                                                                                                                                                                                                           Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Hepatitis C virus (HCV); non-A, non-B hepatitis (HANBH)
                                  Frail DE;
                                                                                                                                                                                                                                                                                                                                                                                                                                           11;
                                                                                                                                                                                                                                                                                                                                                                                                         87.5%; Score 902.5; DB 1
85.8%; Pred. No. 1.1e-83;
Live 9; Mismatches 11
                                  Devare SG,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CTRGVAKAVDEIPVESLETTMRSP 1208
                                                                                                                                                                           Example 1; Page 39-49; 100pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CTRCVAKAVDFIPVESLETTMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        AAP92041 standard; protein; 1766 AA
                                  Desai SM,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          87US-0122714.
87US-0139886.
88US-0161072.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              88EP-0310922
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (updated)
(first entry)
                                                                                                                                                                                                                                                                                                                                                                                                            Query Match
Best Local Similarity 85.8
Matches 175; Conservative
                                  Bode SL, Casey JM, D.
Yamaguchi J, Zeck BJ;
                                                                                                                                                                                                                                                                                                                                                                                3011 AA;
   (ABBO ) ABBOTT LAB
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          18-NOV-1987;
30-DEC-1987;
26-FEB-1988;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     25-MAR-2003
02-MAR-1990
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             18-NOV-1988;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            31-MAY-1989
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              EP318216-A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        AAP92041;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      54
```

RESULT 10

AAP92041 XXX XXX AC AAP XXX XXX XXX DE 141 DE 141 DE 141 DE 315 XXX XXX HEP XXX YXX HEP XXX HEP

Sequence

à g õ g õ 셤 ö

ï

6

348 113 408 ~

```
AAP92047 standard; protein; 2301 AA
                                                                     Disclosure; fig 32; 235pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ပ်
                                                                                                                                                                     87.4%;
85.8%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   87US-0122714.
87US-0139886.
88US-0161072.
88US-0263584.
88US-0263584.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   88EP-0310922
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Kuo
                                                                                                                                                                                                                                                                                                                                                                                                                                           (updated)
(first entry)
                                                                                                                                                                              st Local Similarity 85.8 tches 175; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Houghton M, Choo QL,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  WPI; 1989-159274/22.
N-PSDB; AAN92103.
WPI; 1989-215054/30.
N-PSDB; AAN90331.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CHIR ) CHIRON CORP
                                                                                                                                                    2261 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   18-NOV-1987;
30-DEC-1987;
26-FEB-1988;
06-MAY-1988;
26-OCT-1988;
14-NOV-1988;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   18-NOV-1988;
                                                                                                                                                                                                                                                                                                                                                                                                                                           25-MAR-2003
                                                                                                                                                                                                                                                                                                                                                                                                                                                    02-MAR-1990
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               31-MAY-1989
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           EP318216-A
                                                                                                                                                                                                                                                                                          114
                                                                                                                                                                                                                                                                                                                                                                                                                        AAP92047;
                                                                                                                                                    Sequence
                                                                                                                                                                     Query Match
                                                                                                                                                                                                                                                                                                                                                                                RESULT 13
                                                                                                                                                                                                                                                                                                                                                                                        AAP92047
                                                                                                                                                                                                                                                                                                                                                                                                              DAX PERMENTAL PRINCES SOCIO
                                                                                                                                                                                                              õ
                                                                                                                                                                                                                               d
                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                         δ
                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                               ç
                                                                                                                                                                                                                                                                                                                                               d
                                                                                                                                                                                                                                                                        ::| :::| :::| :::| 289 RRGREILLGPADGMVSKGWRLLAPITAYAQQTRGLLGCIITSLTGRDKNQVEGEVQIVST 348
                                                                                                                                                                                                                                                                                                                                         114 CGSSDLYLVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAV 173
                                                                                                                                                                                                                                                                                                                                                    54 AAQIFLAICINGVCWIVYHGAGIRIIASPKGPVIQMYINVDKDLVGWPAPQGSRSLIPCI 113
                                                                                                                                                                                                                                                                                                             53
                                                                                                                                      The sequence is encoded by the composite cDNA of AAN90327. These antipens react with antibodies in patients with non-A non-B hepatitis (NANBH). They can be used to diagnose HCV-induced NANBH, to raise antibodies for immunoasasy or treatment, or to produce vaccines. (Updated on 25-MAR-2003 to correct PR field.)
                                                                              prodn. of polynucleotide probes, diagnosis, prevention and treatment
                                                                                                                                                                                                                                                              3 KRGSVVIVGRIN------LSGDTAYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVST
                                                                                                                                                                                                                                          9; Gaps
                                                                                                                                                                                                                      DB 10; Length 1786;
                                                                                                                                                                                                                                          Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Hepatitis C virus; clone 12f; clone 15e; probe; vaccine.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Peptide encoded by composite hepatitis C virus cDNA.
                                                                                                                                                                                                                    Query Match

87.4%; Score 901.5; DB 10;
Best Local Similarity 85.8%; Pred. No. 7e-84;
Matches 175; Conservative 9; Mismatches 11;
                                                                                                                                                                                                                                                                                                                                                                                          CTRGVAKAVDFIPVENLETTMRSP 492
                                                                                                                                                                                                                                                                                                                                                                                CTRGVAKAVDFIPVESLETTMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                    AAP90164 standard; protein; 2261 AA
                                                                             Hepatitis C virus gene - used for polypeptide(s) and antibodies for of infection.
                                                                                                                    Disclosure; fig 26; 30pp; English
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Houghton M, Choo QL, Kuo G;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            87US-0122714.
87US-0139886.
88US-0161072.
88US-0263584.
                               Kuo G;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          88GB-0027024
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                (CHIR ) CHIRON CORPORATION.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (first entry)
         (CHIR ) CHIRON CORPORATION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           (updated)
                               Choo QL,
                                                WPI; 1989-215054/30
N-PSDB; AAN90327.
                                                                                                                                                                                                    Sequence 1786 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Pan troglodytes
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      30-DEC-1987;
26-FEB-1988;
26-OCT-1988;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          18-NOV-1988;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              18-NOV-1987;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          25-MAR-2003
01-NOV-1989
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   GB2212511-A
                             Houghton M,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     26-JUL-1989
                                                                                                                                                                                                                                                                                                                                                                                174
                                                                                                                                                                                                                                                                                                                                                                                                                                                                       AAP90164;
                                                                                                                                                                                                                                                                                                                                                                                                                                  RESULT 12
AAP90164
                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                   à
                                                                                                                                                                                                                                                                                                                      a
                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                          셤
                                                                                                                                                                                                                                                                                                                                                                                5
                                                                                                                                                                                                                                                               ö
```

```
439
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         113
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CGSSDLYLVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAV 173
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   23
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Sequence encoded in the hepatitis C virus (HCV) cDNA inserts in clones
12f through 15e.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                54 AAQTFLATCINGVCWTVYHGAGTRTIASPKGPVIQMYTNVDKDLVGWPAPQGSRSLTPCT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               3 KKGSVVIVGRIN------LSGDTAYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVST
                                                                                                                                                              The sequence is the peptide encoded by the composite hepatitis C virus (HCV) CDNA of AAN90331. The polypeptides are used to diagnose HCV-induced NANBH, to raise antibodies for immunoassay or treatment, our to produce vaccines. (Updated on 25-MAR-2003 to correct PR field.)
                                                                                                                                                                                                                                                                                                                                                                                                                                                        9; Gaps
Hepatitis C virus gene - used for prodn. of polynucleotide probes, polypeptide(s) and antibodies for diagnosis, prevention and treatment of infection.
                                                                                                                                                                                                                                                                                                                                                                                             Score 901.5; DB 10; Length 2261;
Pred. No. 9.6e-84;
9; Mismatches 11; Indels 9;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Hepatitis C virus (HCV); non-A, non-B hepatitis (HANBH)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  174 CTRGVAKAVDFIPVESLETTMRSP 197
```

1;

```
It is the sequence encoded in the open reading frame of hepatitis C virus (HCV) cDNA inserts in clones K9-1 through 15e. It is antiqenic and could be used in immunoassay reagents and vaccines and to generate antibodies useful in diagnosis and passive immunotherapy for HCV infection/non-A, (Updated on 25-MAR-2003 to correct PR field.)

(Updated on 25-WAR-2003 to correct PI field.)
                                                                                                                                                                                                                                                                                                                                                                   54 AAQTFLATCINGVCWTVYHGAGTRTIASPKGPVIQMYTNVDKDLVGWPAPQGSRSLTPCT 113
                                                                                                                                                                                                                                                                                                                                                                                                                               114 CGSSDLYLVTRHADVIPVRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAV 173
                                                                                                                                                                                                                                                                                                                                                                                                                                                 3 KKGSVVIVGRIN------LSGDTAYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVST
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Hepatitis C virus gene - used for prodn. of polynucleotide probes, polypeptide(s) and antibodies for diagnosis, prevention and treatment of infection.
                                                                                                                                                                                                                                           Length 2436;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Hepatitis C virus; clone 15e; clone k9-1; probe; vaccine
                                                                                                                                                                                                                                                                       11; Indels
                                                                                                                                                                                                                                       DB 10;
                and associated nucleic acids and polypeptide(s)
                                                                                                                                                                                                                                       87.4%; Score 901.5; DB 1085.8%; Pred. No. 1.1e-83; tive 9; Mismatches 11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Peptide encoded by composite hepatitis C cDNA.
                                                Claim 13; Figure 47-1 - 47-8; 139 pp; English
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           174 CTRGVAKAVDFIPVESLETTWRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Ź
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         AAP90288 standard; protein; 2436
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         87US-0139886.
88US-0161072.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          88GB-0027024
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          87US-0122714
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        88US-0263584
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        (CHIR ) CHIRON CORPORATION.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Kuo
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (first entry)
Purified hepatitis C virus
                                                                                                                                                                                                                                                      Local Similarity 85.8 ses 175; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (updated) (updated)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        choo or,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       WPI; 1989-215054/30.
                                                                                                                                                                                                           2436 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      N-PSDB; AAN90336
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Pan troglodytes
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          18-NOV-1988;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         30-DEC-1987;
26-FEB-1988;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          18-NOV-1987;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     25-MAR-2003
19-JUL-2001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              GB2212511-A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        26-OCT-1988;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      01-NOV-1989
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Houghton M,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            26-JUL-1989
                                                                                                                                                                                                           Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       AAP90288;
                                                                                                                                                                                                                                         Query Match
                                                                                                                                                                                                                                                           Best Loca
Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       셤
                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              δ
     ö
                                                                                                                                                                                                                                                                                                                                                                                                                               õ
                                                                                                                                                                                                                                                                                                                                     It is the sequence encoded in the open reading frame of hepatitis C virus (MCV) cDNA inserts in clones 12f through 15e. It is antigenic and could be used in immunoassay reagents and vaccines and to generate antibodies useful in diagnosis and passive immunotherapy for HCV infection/non-A,
                                                                                                                                                                                                                                                                                                                                                                                  AAQTFLATCINGVCWTVYHGAGTRTIASPKGPVIQMYTNVDKDLVGWPAPQGSRSLTPCT 113
                                                                                                                                                                                                                                                                                                                                                                                                                                              114 CGSSDLYLVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAV 173
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Sequence encoded in the hepatitis C virus (HCV) cDNA inserts in clones K9-1 through 15e.
                                                                                                                                                                                                                                                                                                                    3 KKGSVVIVGRIN-----LSGDTAYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVST
                                                                                                                                                                                                                                                                                        Gaps
                                                                                                                                                                                                                                                      Length 2301;
                                                                                                                                                                                                                                                                                        6
                                                                                                                                                                                                                                                      87.4%; Score 901.5; DB 10; Length
85.8%; Pred. No. 9.9c-84;
ive 9; Mismatches 11; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Hepatitis C virus (HCV); non-A, non-B hepatitis (HANBH)
                                   - and associated nucleic acids and polypeptide(s)
                                                                                                                                                           non-B hepatitis.
(Updated on 25-MAR-2003 to correct PR field.)
(Updated on 25-MAR-2003 to correct PI field.)
                                                                 Claim 13; Figure 32-1 - 32-7; 139 pp; English
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             174 CTRGVAKAVDFIPVESLETIMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CIRGVAKAVDFIPVENLETIMRSP 583
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Ź
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        AAP92050 standard; protein; 2436
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        87US-0139886.
88US-0161072.
88US-0191263.
88US-0263584.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           88EP-0310922
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          87US-0122714
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       88US-0271450
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Kuo
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (first entry)
                  Purified hepatitis C virus
                                                                                                                                                                                                                                                                                        Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       (updated)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Houghton M, Choo QL,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      WPI; 1989-159274/22
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       (CHIR ) CHIRON CORP
                                                                                                                                                                                                                          2301 AA
                                                                                                                                                                                                                                                                     Local Similarity
es 175; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       N-PSDB; AAN92106
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           18-NOV-1988;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         18-NOV-1987;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    02-MAR-1990
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        26-FEB-1988;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         06-MAY-1988;
26-0CT-1988;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     25-MAR-2003
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     14-NOV-1988
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           31-MAY-1989
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           30-DEC-1987
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             EP318216-A.
```

54

셤 ò a á g ó 셤

Sequence Query Match Best Loc Matches

200

AAP92050;

AAP92050
XX
AC
AAP9
XX
AAP9
DT
CS-k
DT

RESULT 14

614

ŝ

```
ij
                                                                                                                                                                                                                                54 AAQIFLATCINGVCWIVYHGAGIRIIASPKGPVIQMYINVDKDLVGWPAPQGSRSLIPCT 113
                                                                                                                                                                                                                                                                                                          114 CGSSDLYLVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAV 173
                                                                                                                                                                                                                                                                                                                                                                   The sequence is the peptide encoded by the composite hepatitis C virus (HCV) cDNA of AAN90336. The polypeptides are used to diagnose HCV-induced NANBH, to raise antibodies for immunoassay or treatment, or to produce vaccines.

(N.B. This record was resubmitted to correct errors in the sequence.) (Updated on 25-MAR-2003 to correct PR field.)
                                                                                                                                                                                                      9; Gaps
                                                                                                                                                                      Query Match

87.4%; Score 901.5; DB 10; Length 2436;
Best Local Similarity 85.8%; Pred. No. 1.1e-83;
Matches 175; Conservative 9; Mismatches 11; Indels 9;
            Disclosure; fig 47-1 to 47-8; 235pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                              174 CTRGVAKAVDFIPVESLETTMRSP 197
                                                                                                                                             2436 AA;
                                                                                                                                               Sequence
õ
                                                                                                                                                                                                                                                           QQ
                                                                                                                                                                                                                                                                                          õ
                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                    QQ
```

Search completed: August 30, 2003, 19:12:23 Job time : 45.6227 secs

Н

```
Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ô
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        genome polyprotein
polyprotein - marm
hypotheticai prote
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               angio-associated m
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    hypothetical prote
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                nitrate/nitrite se
probable aromatic
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            beta transducin ho
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                transferrin-bindin
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       probable periplasm
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   hypothetical prote
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             hypothetical prote
probable exported
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         hypothetical prote
ferredoxin-nitrite
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            polyprotein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        polyprotein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         nitrate/nitrite se
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             polyketide synthas
                                                                                                                                                               (without alignments)
1168.492 Million cell updates/sec
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.
                                                                                                                                                                                                                                                                    1 MKKKGSVVIVGRINLSGDTA.....VAKAVDFIPVESLETTMRSP 197
                                                                                                                                     August 30, 2003, 19:02:22 ; Search time 16.2134 Seconds
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Description
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   genome
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        genome
                                                                                                                                                                                                                                                                                                                                                                                                                            283308
  5.1.6
Compugen Ltd.
                                                                                                                                                                                                                                                                                                                                                                                                                          hits satisfying chosen parameters:
                                                                                                                                                                                                                                                                                                                                                                                   283308 seqs, 96168682 residues
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SUMMARIES
GenCore version
Copyright (c) 1993 - 2003
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

    protein search, using sw model

                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 T08839
B71360
I39383
E75392
B81104
C81911
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 GNWVC3
S40770
GNWVCH
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               GNWVTW
GNWVTC
GNWVCJ
A45573
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                $18030
JC5620
JQ1303
GNWVJ8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            F81196
B71284
S28261
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             T06118
AH1030
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            08841
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   r49806
                                                                                                                                                                                                                                                                                                                                     Gapop 10.0 , Gapext 0.5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                        length: 0
length: 2000000000
                                                                                                                                                                                                                             US-09-965-594-14
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            %
Query
Match Length DB
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               piri: *
pir2: *
pir3: *
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                PIR_76:*
                                                                                                                                                                                                                                                                                                                  BLOSUM62
                                                                                                                                                                                                                                                                                                                                                                                                                               Fotal number of
                                                                                                                                                                                                                                                                                                                                                                                                                                                                        DB seq
DB seq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             76.5
76.5
76
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Score
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               9901.5
9900.5
9900.5
8886.5
8886.5
8860.5
8846.5
6746.5
6746.5
8766.5
8766.5
8766.5
8766.5
8766.5
8766.5
8766.5
8766.5
8766.5
8766.5
8766.5
8766.5
8766.5
8766.5
8766.5
8766.5
8766.5
8766.5
8766.5
8766.5
8766.5
8766.5
8766.5
8766.5
8766.5
8766.5
8766.5
8766.5
8766.5
8766.5
8766.5
8766.5
8766.5
8766.5
8766.5
8766.5
8766.5
8766.5
8766.5
8766.5
8766.5
8766.5
8766.5
8766.5
8766.5
8766.5
8766.5
8766.5
8766.5
8766.5
8766.5
8766.5
8766.5
8766.5
8766.5
8766.5
8766.5
8766.5
8766.5
8766.5
8766.5
8766.5
8766.5
8766.5
8766.5
8766.5
8766.5
8766.5
8766.5
8766.5
8766.5
8766.5
8766.5
8766.5
8766.5
8766.5
8766.5
8766.5
8766.5
8766.5
8766.5
8766.5
8766.5
8766.5
8766.5
8766.5
8766.5
8766.5
8766.5
8766.5
8766.5
8766.5
8766.5
8766.5
8766.5
8766.5
8766.5
8766.5
8766.5
8766.5
8766.5
8766.5
8766.5
8766.5
8766.5
8766.5
8766.5
8766.5
8766.5
8766.5
8766.5
8766.5
8766.5
8766.5
8766.5
8766.5
8766.5
8766.5
8766.5
8766.5
8766.5
8766.5
8766.5
8766.5
8766.5
8766.5
8766.5
8766.5
8766.5
8766.5
8766.5
8766.5
8766.5
8766.5
8766.5
8766.5
8766.5
8766.5
8766.5
8766.5
8766.5
8766.5
8766.5
8766.5
8766.5
8766.5
8766.5
8766.5
8766.5
8766.5
8766.5
8766.5
8766.5
8766.5
8766.5
8766.5
8766.5
8766.5
8766.5
8766.5
8766.5
8766.5
8766.5
8766.5
8766.5
8766.5
8766.5
8766.5
8766.5
8766.5
8766.5
8766.5
8766.5
8766.5
8766.5
8766.5
8766.5
8766.5
8766.5
8766.5
8766.5
8766.5
8766.5
8766.5
8766.5
8766.5
8766.5
8766.5
8766.5
8766.5
8766.5
8766.5
8766.5
8766.5
8766.5
8766.5
8766.5
8766.5
8766.5
8766.5
8766.5
8766.5
8766.5
8766.5
8766.5
8766.5
8766.5
8766.5
8766.5
8766.5
8766.5
8766.5
8766.5
8766.5
8766.5
8766.5
8766.5
8766.5
8766.5
8766.5
8766.5
8766.5
8766.5
8766.5
8766.5
8766.5
8766.5
8766.5
8766.5
8766.5
8766.5
8766.5
8766.5
8766.5
8766.5
8766.5
8766.5
8766.5
8766.5
8766.5
8766.5
8766.5
8766.5
8766.5
8766.5
8766.5
8766.5
8766.5
8766.5
8766.5
8766.5
8766.5
8766.5
8766.5
8766.5
8766.5
8766.5
8766.5
8766.5
8766.5
8766.5
8766.5
8766.5
8766.5
8766.5
8766.5
8766.5
8766.5
8766.5
8766.5
8766.5
8766.5
8766.5
8766.5
8766.5
8766.5
8766.5
8766.5
8766.5
8766.5
8766.5
8766.5
8766.5
87
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      80.5
80.5
78.5
78.5
77.5
                                                                                                                                                                                                                                                  Perfect score:
                                                                                                                                                                                                                                                                                                                     Scoring table:
                                                                                              OM protein
                                                                                                                                                                                                                                                                           Sequence:
                                                                                                                                                                                                                                                                                                                                                                                     Searched:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Database
                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Minimum
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Maximum
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           78 7 7 7 7 7 7 7 7 8 7 8 7 8 7 8 7 8 8 7 8 8 7 8 8 7 8 8 7 8 8 7 8 8 7 8 8 7 8 8 7 8 8 7 8 8 7 8 8 7 8 8 8 7 8
                                                                                                                                        Run on:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Result
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ĕ.
```

transferrin-bindin surface-associated	probable regulator hypothetical prote	glycosyl transfera	transferrin-bindin	orsellinic acid sy	zona pellucida gly	hypothetical prote	hypothetical prote	hypothetical prote	ferredoxin-nitrite	HIV-1 retropepsin	conserved hypothet	hypothetical prote
JN0821 S61441	C96033 AB1775	D87603	C81832	T30871	S70401	AH2406	S23441	T04533	S30922	GNLJGL	G87265	876618
77	77	٦,	1 0	7	~	~	7	N	~	٦	~	7
911 2508	329 1334	322	910	1293	415	424	755	846	456	1165	239	317
7.4	7.3	7.3	7.3	7.3	7.2	7.2	7.2	7.2	7.2	7.2	7.1	7.1
76	75.5	75	75	75	74.5	74.5	74.5	74.5	74	74	73.5	73.5
3.0	33	34	36	37	38	39	40	41	42	43	44	45

## ALIGNMENTS

```
A; Status: preliminary
A; Residues: 1377-1633 -CH2>
C; Superfamily: hepatitis C virus genome polyprotein
C; Superfamily: hepatitis C virus genome polyprotein; glycoprotein; hydrolase; nonstruc
C; Superfamily: hepatitis C virus genome polyprotein
C; Superfamily: hepatitis C virus genome protein R ** status predicted -CPE>
F; 116-131/Product: monstructural protein NS1 ** status predicted -NS1>
F; 130-1237/Region: nucleotide-binding motif A (P-loop)
F; 1312-1317/Region: nucleotide-binding motif B
F; 1312-1317/Region: nucleotide-binding motif B
F; 1316-1862/Product: nonstructural protein NS4* ** status predicted -N4A>
F; 2014-3011/Product: nonstructural protein NS4* ** status predicted -NS5>
F; 2014-3011/Product: nonstructural protein NS4* ** status predicted -NS5>
F; 196, 209, 234, 305, 325, 417, 423, 430, 448, 476, 532, 540, 556, 576, 623, 645, 1213, 1207
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        A.Accession: A39166
A.Molecule type: mRNA
A.Rocassidues: 1-3011 < CHO>
A.Rocassidues: 1-3011 < CHO>
A.Rocassidues: 1-3012 < CHO>
A.Rocassidues: 1-3012 < CHO>
A.Rocassidues: 1-3012 < CHO>
A.Rocassidues: 1-3012 < CHO>
A.COCASS references: GB:M62321; NID:g329873; PIDN:AAA45676.1; PID:g329874
A.Cocassidues: 1-3012 < CHO-
B.C. Mailysis of a new hepatitis C virus type and its phylogenetic relationship A.Reference number: PQ0393; MUID:g2268871; PMID:1316939
A.Rocassidues: 1970-1633 < CHA>
A.Rocassidues: 1577-1633 < CHA>
A.Rocassidues: 1577-1633 < CHA>
A.Cocassidues: 1577-1673 < CHA>
A.Cocassidues: 1577-167
genome polyprotein - hepatitis C virus (strain HCV-1)
N;Contains: capsid protein C; envelope protein M; hepacivirin (EC 3.4.21.98) (nonstr
protein NS4a; nonstructural protein NS4b; nonstructural protein NS5
                                                                                                                                                                                                         C;Species: hepatitis C virus
C;Date: 30-Sep-1992 *Sequence_revision 30-Sep-1992 *text_change 19-Jan-2001
C;Date: 30-Sep-1992 *Sequence_revision 30-Sep-1992 *text_change 19-Jan-2001
C;Accession: A99166; PQQA03; PQQA048
R;Choo, O.L.; R;Chman, K.H.; Han, J.H.; Berger, K.; Lee, C.; Dong, C.; Gallegos, C.; Prico, O.L.; R;Chman, K.H.; Han, J.H.; Berger, K.; Lee, C.; Dong, C.; Gallegos, C.; Prico, Natl. Acad. Sci. U.S.A. 88, 2451-2455, 1991
A;Title: Genetic organization and diversity of the hepatitis C virus.
A;Reference number: A39166; MUID:91172826; PMID:1848704
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               53
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       3 KKGSVVIVGRIN-----LSGDTAYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVST
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  87.4%; Score 901.5; DB 1; Length 3011; 85.8%; Pred. No. 3.7e-75;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            11; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    9; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Matches 175; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Local Similarity
```

```
R; Inchauspe, G; Zebedee, S; Lee, D.H.; Sugitani, M; Nasoff, M.; Prince, A.M. Proc. Natl. Acad. Sci. U.SA. 88, 10292-10296, 1991

R; Inchauspe, G; Zebedee, S; Lee, D.H.; Sugitani, M; Nasoff, M; Prince, A.M. Proc. Natl. Acad. Sci. U.SA. 88, 10292-10296, 1991

A; Tille; Ganomic structure of the human prototype strain H of hepatitis C virus: comp A; Reference number: A41546; MUID:92052256; PMID:1658800

A; Contents: annotation

A; Contents: annotation

A; Contents: annotation

C; Superfamily: hepatitis C virus genome polyprotein

C; Superfamily: hepatitis C virus genome polyprotein

C; Superfamily: hepatitis C virus genome polyprotein

C; Reywords: ATP; capsid protein; envelope protein; glycoprotein; hydrolase; nonstruct

F; 115/Product: envelope protein E $status predicted <NSD-F; 190-129/Region: nonstructural protein NSI $status predicted <NSD-F; 1007-1615/Product: nonstructural protein NSI $status predicted <NSD-F; 11007-1615/Product: nonstructural protein NSI $status predicted <NSD-F; 11007-1615/Product: nonstructural protein more protein of $1000 |

F; 1130-1237/Region: nucleotide-binding motif A (P-loop)

F; 1131-1317/Region: nucleotide-binding motif E (P-loop)
                                                                                                     M.Contains: capsid protein - hepatitis C virus (strain H)
N.Contains: capsid protein C; envelope protein H; hepacivirin (EC 3.4.21.9B) (nonstruprotein NS4s; nonstructural protein NS4s; nonstructural protein NS5
C;Species: hepatitis C virus
A;Note: host Homo sapiens (man)
C;Date: 31-Dec-1992 #sequence_revision 31-Dec-1992 #text_change 19-Jan-2001
C;Accession: A36814; A41546
R;Inchauspe, G; Zebedee, S;;Lee, D.H.; Sugitani, M.; Nasoff, M.; Prince, A.M.
S;Unchauspe, G; Zebedee, S;;Lee, D.H.; Sugitani, M.; Nasoff, M.; Prince, A.M.
A;Description: Genomic structure of the human prototype strain H of hepatitis C virus
A;Accession: A36814
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           A;Molecule type: genomic RNA
A;Residues: 1-3011 <INC>
A;Cross-references: GB:M67463; NID:9329737; PIDN:AAA45534.1; PID:9329738
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Query Match
Best Local Similarity
Matches 171; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     GNWVTW
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  A.Experimental source: isolate HC-JI

C.Superfamily: hepatitis C virus genome polyprotein

C.Keywords: ATP; glycoprotein; hydrolase; nucleotide binding; P-loop; polyprotein; serin

C.Keywords: ATP; glycoprotein; hydrolase; nucleotide binding; P-loop; polyprotein; serin

F; 115-115/Product: envelope protein C *status predicted <CPC>

F; 116-191/Product: major envelope protein B *status predicted <NEE>

F; 390-729/Product: major envelope protein NSI *status predicted <NSI>

F; 300-1006/Product: nonstructural protein NSI *status predicted <NSI>

F; 1230-1237/Region: nucleotide-binding motif A (P-loop)

F; 1312-1317/Region: DEXH motif
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       A;Cross_references: EMBL:D10749; NID:9221586; PIDN:BAA01582.1; PID:9221587
R;Okamoto, H.; Okada, S.; Sugiyama, Y.; Yotsumoto, S.; Tanaka, T.; Yoshizawa, H.; Tsuda, Jpn. J. Exp. Wed. 60, 167-177, 1990
A;Title: The S'-terminal sequence of the hepatitis C virus genome.
A;Reference number: PC1284; MUID:91013116; PMID:2170712
A;Accession: PC1285
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    protein M; hepacivirin (EC 3.4.21.98) (nonstructu
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1005 RRGREILLGPANGMVSKGWRLLAPITAYAQQTRGLLGCIITSLTGRDKNQVEGEVQIVST 1064
                                                                                                     113
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       114 CGSSDLYLVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAV 173
                                                                                                                                                                                                114 CGSSDLYLVTRHADVIPVRRRGDSRGSLLSPRPISYI.KGSSGGPLLCPAGHAVGIFRAAV 173
                                                                  AAOTFLATCINGVCWTVYHGAGTRTIASPKGPVIQMYTNVDKDLVGWPAPQGSRSLTPCT 113
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   genome polyprotein - hepatitis C virus
N;Contains: capsid protein C; envelope protein M; hepacivirin (EC 3.4.21.98)
protein NS4s: nonstructural protein NS4b; nonstructural protein NS5
C;Species: hepatitis C virus
C;Species: hepatitis C virus
C;Date: 19-May-2000 *sequence_revision 19-May-2000 *text_change 19-Jan-2001
C;Accession: S40770; PC1285
R;Okamoto, H.
A;Reference number: S40770
A;Reference number: S40770
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    3 KKGSVVIVGRIN-----LSGDTAYAQQTRGEEGCQETSQTGROKNQVEGEVQIVST
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     AAQTFLATCINGVCWTVYHGAGTRTIASPKGPVIQMYTNVDKDLVGWPAPQGSRSLTPCT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           A;MOlecule type: genomic RNA
A;Residues: 1.513 <OK2>
A;Cross-references: GB:D00831; NID:g221511; PIDN:BAA00705.1; PID:g221512
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              F;1616-1862/Product: nonstructural protein NS4a #status predicted <N4A>F;1863-2013/Product: nonstructural protein NS4b #status predicted <N4B>F;2014-3011/Product: nonstructural protein NS5 #status predicted <NS5>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ..
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Length 3011;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        87.3%; Score 900.5; DB 1;
85.8%; Pred. No. 4.5e-75;
rative 8; Mismatches 12;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              174 CIRGVAKAVDFIPVESLETIMRSP 197
                                                                                                                                                                                                                                                                                                                                     174 CTRGVAKAVDFIPVESLETTMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   A:Molecule type: genomic RNA A:Residues: 1-3011 < OKA>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Local Similarity
hes 175; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            õ
                                                                                                                          g
                                                                                                                                                                                                   õ
                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                         g
```

```
genome polyprotein - hepatitis C virus (strain Taiwan)
N;Contains: capsid protein C; envelope protein M; hepacivirin (EC 3.4.21.98) (nonstruu protein NS4s, nonstructural protein NS5 c; Species: hepatitis C virus
A;Note: host Homo sapiens (man)
C;Date: host Homo sapiens (man)
C;Date: al-Dec-1992 #sequence_revision 31-Dec-1992 #text_change 19-Jan-2001
C;Accession: A40244
R;Chen, P.J; Lin, M.H.; Tai, K.F.; Liu, P.C.; Lin, C.J.; Chen, D.S.
Virology 188, 102-113, 1992
F;1616-1862/Product: nonstructural protein NS4a *status predicted <NV4>
F;1863-2013-Product: nonstructural protein NS4b *status predicted <NV8>
F;2014-3011/Product: nonstructural protein NS5 *status predicted <NS5>
F;2014-3011/Product: nonstructural protein NS5 *status predicted <NS5>
F;196,209,234,305,325,417,423,430,448,476,532,540,556,576,623,645,1213,1255,2041,2240
                                                                                                                                                                                                                                                                                                                                           54 AAQTFLATCINGVCWTVYHGAGTRTIASPKGPVIQMYTNVDKDLVGWPAPQGSRSLTPCT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  3 KKGSVVIVGRIN-----LSGDTAYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVST
                                                                                                                                                                                                                                      Gaps
                                                                                                                                                                                                                                  6
                                                                                                                                                                Length 3011;
                                                                                                                                                                                                                                  Indels
                                                                                                                                                                85.9%; Score 886.5; DB 1; 83.8%; Pred. No. 9.2e-74;
                                                                                                                                                                                                                                  Conservative 10; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    174 CTRGVAKAVDFIPVESLETTMRSP 197
```

Gaps

Length 3010;

```
A; McDecule type: genomic RNA
A; Residues: 2650-2707 < KA2>
A; Residues: 2550-3707 < KA2
A; Residues: 2550-3707 < KA2
C; Comment: The cleavage sites of this polyprotein have not been determined.
C; Comment: The cleavage sites of this polyprotein
C; Superfamily: hepatitis C virus genome polyprotein
F; 2115/Product: capsid protein M #status predicted < KES>
F; 300-1006/Product: nonstructural protein NSI #status predicted < NS2>
F; 300-1006/Product: nonstructural protein NSI #status predicted < NS2>
F; 1230-1237/Region: nucleotide-binding motif A (P-loop)
F; 1312-1317/Region: nucleotide-binding motif A (P-loop)
F; 1312-1317/Region: nucleotide-binding motif A (P-loop)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            R.Kato, N.; Hijikata, M.; Ootsuyama, Y.; Nakagawa, M.; Ohkoshi, S.; Sugimura, T.; Sh Proc. Natl. Acad. Sci. U.S.A. 87, 9524-9528, 1990
A.711tle: Molecular cloning of the human hepatitis C virus genome from Japanese patie A.71ttle: Molecular cloning of the human hepatitis C virus genome from Japanese patie A.74cession. A39253
A.Reference number: A39253; MUID:91088550; PMID:2175903
A.Reference number: A39253
A.Reference number: B30208; NID:9221610; PIDN:BAA14233.1; PID:9221611
R.Kato, N.; Ohkoshi, S.; Shimotohno, K.
Proc. Jpn. Acad. 65B, 219-223, 1989
A.71tle: Japanese isolates of the non-A, non-B hepatitis viral genome show sequence A.Reference number: PS0086
A.Reference number: PS0086
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  F:1616-1862/Product: nonstructural protein NS4a *status predicted <N4A>
F:1863-2013/Product: nonstructural protein NS4b *status predicted <N4B>
F:2014-3010/Product: nonstructural protein NS5 *status predicted <NS5>
F:205,209,234,250,305,325,417,423,430,448,532,556,576,623,645,1213,1255,2041,2077,22
F;196,209,234,250,305,325,417,423,430,448,532,540,556,576,623,645,1213,1255,2041,207
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1005 RRGKEILLGPADSFGEQGWRLLAPITAYSQQTRGLLGCIITSLTGRDKNQVDGEVQVLST 1064
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1125 CGSSDLYLVTRHADVIPVRRRGDSRGSLLSPRPVSYLKGSSGGPLLCPFGHAVGIFRAAV 1184
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1065 ATQSFLATCVNGVCWTVYHGAGSKTLAAPKGPITQMYTNVDQDLVGWPKPPGARSLTPCT 1124
                                                                                                                                                                                                                                                                                                             114 CGSSDLYLVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAV 173
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     genome polyprotein - hepatitis C virus (strain J)
N;Contains: capsid protein C; envelope protein M; major envelope protein E;
protein NS4a: nonstructural protein NS4b; nonstructural protein NS5
C; Species: hepatitis C virus
C; Species: hepatitis C virus
C; Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 19-Jan-2001
C;Accession: A39253; ps0086
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    54 AAQTFLATCINGVCWTVYHGAGTRTIASPKGPVIQMYTNVDKDLVGWPAPQGSRSLTPCT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              3 KRGSVVIVGRINLSGD-----TAYAQQTRGEEGCQETSQTGRDKWQVEGEVQIVST
                                                                                                                                                                                                                                                                                                                                                                                                                                      AAQTFLATCINGVCWTVYHGAGTRTIASPKGPVIQMYTNVDKDLVGWPAPQGSRSLTPCT
                                                                                                                                                                                                                                                                3 KKGSVVIVGRIN------LSGDTAYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVST
                                                                            Score 850.5; DB 1;
Pred. No. 2.1e-70;
L: Mismatches 16;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Score 850.5; DB 1 Pred. No. 2.1e-70;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   23; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           174 CTRGVAKAVDFIPVESLETIMRSP 197
                                                                                                                                                                 21;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              82.4%;
                                                                                82.4%;
illarity 77.5%;
Conservative 21
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    156; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Local Similarity
                                                                                     Query Match
Best Local Similarity
Matches 158; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                      54
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1185
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             RESULT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          GNWVCJ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ö
                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                    ô
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              C; Species: hepatitis C virus
C; Species: hepatitis C virus
C; Date: 31-Mar-1992 $sequence_revision 31-Mar-1992 $text_change 19-Jan-2001
A; Takamisawa, A.; Mori, C.; Fuke, I.; Manabe, S.; Murakami, S.; Fujita, J.; Onishi, E.;
J. Virol. 65, 1105-1113, 1991
A; Teference number: A38465; MuID:91140698; PMID:1847440
A; Residues: 1-3010 < TAK.
A; Residues: 1-3010 < TAK.
A; Residues: 1-3010 < TAK.
A; Cross-references: EMBL:M58335; NID:9329770; PIDN:AAA72945.1; PID:9329771
C; Superfamily: hepatitis C virus genome polyprotein
C; Superfamily: Hepatitis C virus Genome isolated from human
C; Superfamily: Hepatitis C virus Genome 
                                                                                                                                                                                            A.Cross-references: GB:M84754
C:Superfamily: hepatitis C virus genome polyprotein
C:Superfamily: hepatitis C virus genome polyprotein
C:Superfamily: hepatitis C virus genome polyprotein; hydrolase; nonstructura
C:Keywords: ATP; capsid protein C *status predicted <CPC>
F:115/Product: capsid protein C *status predicted <CPC>
F:116-191/Product: major envelope protein M *status predicted <MEE>
F:192-389/Product: major envelope protein NSI *status predicted <NSI>
F:190-106/Product: nonstructural protein NSI *status predicted <NSI>
F:120-1137/Region: nucleotide-binding motif A (P-loop)
F:1312-1317/Region: DEXH motif
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       genome polyprotein - hepatitis C virus
N;Contains: capsid protein C; envelope protein M; hepacivirin (EC 3.4.21.98) (nonstructu
protein NS4a; nonstructural protein NS4b; nonstructural protein NS5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      F;1616-1862/Product: nonstructural protein NS4a #status predicted <N4A>
F;1616-2013/Product: nonstructural protein NS4b #status predicted <N4B>
F;2014-3010/Product: nonstructural protein NS5 #status predicted <NS5>
F;2014-3010/Product: nonstructural protein NS5 #status predicted <NS5>
F;196,209,233,234,250,305,325,417,423,430,448,532,540,556,576,623,645,1213,1255,2041,207
A:Title: The Taiwanese hepatitis C virus genome: sequence determination and mapping the A:Reference number: A40244; MUID:92230206; PMID:1314449
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             54 AAQTFLATCINGVCWTVYHGAGTRTIASPKGPVIQMYTNVDKDLVGWPAPQGSRSLTPCT 113
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            114 CGSSDLYLVTRHADVIPVRRRCDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAV 173
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           53
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           3 KKGSVVIVGRIN-----LSGDTAYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVST
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          F;1616-1862/Product: nonstructural protein NS4a #status predicted <N4A>F;1863-2013/Product: nonstructural protein NS4b #status predicted <N4B>F;2014-3010/Product: nonstructural protein NS5 #status predicted <NS5>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          9.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Length 3010;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          16; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     82.8%; Score 854.5; DB 1; 78.9%; Pred. No. 8.7e-71;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Ouery Match 82.8%; Score 854.5;
Best Local Similarity 78.9%; Pred. No. 8.7e
Matches 161; Conservative 18; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CTRGVAKAVDFIPVESLETTMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  F;1316-1319/Region: DEXH motif
                                                                                                                                  A; Molecule type: genomic RNA
A; Residues: 1-3010 <CHE>
                                                                                                  A; Accession: A40244
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     đ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   à
```

protein E; NS5

6

Indels

Length 3010;

```
C;Date: 19-May-2000 #sequence_revision 19-May-2000 #text_change 23-Mar-2001
C;Accession: S18030; S33570; A48332; S18029
R;Honda, M.; Kaneko, S.; Masashi, U.; Kobayashi, K.; Murakami, S.
submitted to the EMBL Data Library, September 1991
A;Description: A whole genome of hepatitis C virus cDNA was isolated from a single pa A;Accession: S18030
                                                                                                                                                                                                                                                                                                                   A; Experimental source: isolate JK1 from an individual
R; Honda, M.; Kaneko, S.; Unoura, M.; Kobayashi, K.; Murakami, S.
Arch. Virol. 128, 163-169, 1993
A; Title: Sequence analysis of putative structural regions of hepatitis C virus isolat
A; Reference number: A48332; MUID:93119270; PMID:8380322
A; Accession: S33570
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     F;1616-1862/Product: nonstructural protein NS4a *status predicted <N4A>
F;1863-2013/Product: nonstructural protein NS4 *status predicted <N4B>
F;2014-3010/Product: nonstructural protein NS5 *status predicted <NS5>
F;2014-3010/Product: nonstructural protein NS5 *status predicted <NS5>
F;196,209,234,250,305,417,423,448,532,540,556,576,623,645/Binding site: carbohydrate
                                                                                                                                                                                                                               A;Wolecule type: genomic RNA
A;Residues: 1-3010 <HON>
A;Cross:references: EMBL:X61596; NID:g59478; PIDN:CAA43793.1; PID:g59479
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Best Local Similarity 76.5%;
Matches 156; Conservative 2
        A; Variety: isolate JK1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ô
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       A;Status: preliminary
A;Status: preliminary
A;Status: preliminary
A;Status: preliminary
A;Rosledues: 1-3010 cr1AN>
A;Rosledues: 1-3010 cr1AN>
A;Cross-references: GB:D11168; GB:D01171; NID:g221612; PIDN:BAA01943.1; PID:g221613
A;Cross-reference extracted from NCBI backbone (NCBIN:106206, NCBIP:106207)
C;Superfamily: hepatitis C virus genome polyprotein planting; P-1000; polyprotein; Serin C;Superfamily: hepatitis C virus genome polyproteide binding; P-1000; polyprotein; SF:115-79roduct: capsid protein C *status predicted cERN>
F;116-191/Product: envelope protein M *status predicted cNED>
F;116-191/Product: nonstructural protein NS1 *status predicted cNS2>
F;130-1066/Product: nonstructural protein NS2 *status predicted cNS2>
F;1207-1217/Region: nucleoctide-binding motif A (P-100p)
F;1312-1317/Region: bEXH motif
F;1616-1862/Product: nonstructural protein NS4 *status predicted cN4A>
F;1616-1862/Product: nonstructural protein NS4 *status predicted cN45>
F;1014-3010/Product: nonstructural protein NS4 *status predicted cN45>
F;1014-3010/Product: nonstructural protein NS5 *status predicted cN55>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       genome polyprotein - hepatitis C virus (isolate JK1)
N;Contains: capsid protein C; envelope protein M; hepacivirin (EC 3.4.21.98) (nonstructup protein NS4a; nonstructural protein NS4b; nonstructural protein NS5 C;Species: hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                   genome polyprotein - hepatitis C virus (strain JT)
N;Contains: capsid protein C; envelope protein M; hepacivirin (EC 3.4.21.98) (nonstructu
protein NS4a; nonstructural protein NS4b; nonstructural protein NS5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 R;Tanaka, T.; Kato, N.; Nakagawa, M.; Ootsuyama, Y.; Cho, M.J.; Nakazawa, T.: Hijikata, Virus Res. 23, 39-53, 1992
A;Title: Molecular cloning of hepatitis C virus genome from a single Japanese carrier: A;Reference number: A45573; MoID:92295714; PMID:1318627
A;Status: preliminary
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1065 ATGSFLATCVNGVCWTVYHGAGSKTLAGPKGPITQMYINVDQDLVGWPAPPGARSWTPCT 1124
                                                                                                                       54 AAQIFLATCINGVCWTVYHGAGIRTIASPKGPVIQMYINVDKDLVGWPAPQGSRSLIPCT 113
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             114 CGSSDLYLVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAV 173
                                                                                                                                                                                                                                                                                                                                                                                                                                                                              C;Species: hepatitis C virus
C;Date: 19-May-2000 *sequence_revision 19-May-2000 *text_change 19-Jan-2001
C;Accession: A45573
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       3 KRGSVVIVGRIN-----LSGDTAYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVST
                                                                                      CGSSDLYLVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 9; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        DB 1; Length 3010;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 17; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     tch 81.8%; Score 844.5; DB 1 al Similarity 77.5%; Pred. No. 7.4e-70; 158; Conservative 20; Mismatches 17
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CIRCVAKAVDFIPVESMETTMRSP 1208
                                                                                                                                                                                                                               1185 CTRGVAKAVDFIPVESMETTMRSP 1208
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               174 CTRGVAKAVDFIPVESLETTMRSP 197
                                                                                                                                                                                                 174 CTRGVAKAVDFIPVESLETIMRSP 197
                                                                                                                                            1125
                                                                                      114
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Best Local
Matches 15
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           RESULT 8
S18030
                                                                                                                                                                                                                                                                                                                                             RESULT 7
                                                                                                                                                                                                 à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                               g
                                                                                      ò
```

se

```
Richamberlain, R.W.; Adams, N.J.; Taylor, L.A.; Simmonds, P.; Elliott, R.M.
Blochem, Blochbys, Res. Commun. 236, 44-49, 1997
A;Title: The complete coding sequence of hepatitis C virus genotype 5a, the predomina A;Reference number: JC5620; MUID:97366593; PMID:9223423
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        genome polyprotein - hepatitis C virus (isolate EUH1480)
N:Contains: capsid protein C; envelope protein M; hepacivirin (EC 3.4.21.98) (nonstru-
protein NS4a; nonstructural protein NS4b; nonstructural protein NS5
                                                                                                                                                                                                                                                                                                                                        1124
                                                                                                                                                                                                          1064
                                                                                                                                                                                                                                                                                                                                                                                                                                       C.Species: hepatitis C virus
C.Date: 19-May-2000 #sequence_revision 19-May-2000 #text_change 19-Jan-2001
C.Accession: JC5620
                                                                                                                                                                           54 AAQIFLAICINGVCWTVYHGAGTRIIASPKGPVIOMYTNVDKDLVGWPAPQGSRSLIPCT
                                                                                                                                                                                                                                                                                                                                        1065 ATOSFLATCVNGVCWTVYHGAGSKTLAGPKGPINOMYTNVDQDLVGWQAPSGAASLTPCT
                                                                                                                                                                                                                                                                                                                                                                                                    114 CGSSDLYLVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAV
                                                                                                                                    3 KKGSVVIVGRIN-----LSGDTAYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVST
                                                                            Gaps
80.3%; Score 828.5; DB 1; Length 3010; 76.5%; Pred. No. 2.3e-68;
                                                                     Indels
                                                                  20; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        174 CTRGVAKAVDFIPVESLETTMRSP 197
```

```
F;2018-3033/Product: nonstructural protein NS5 #status predicted <NO5>
F;196,209,234,305,325,417,423,430,448,477,534,542,558,578,627,649,1091,1217,1259,2036
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Matches 123; Conservative
                                                                                                                                                                                                                                                Matches 122; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         A; Molecule type: mRNA
A; Residues: 2678-2729 <KAT>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Query Match
Best Local Similarity
                                                                                                                                                         Query Match
Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                             q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ద
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        C; Accession: J01303
R; Okamoto, H.; Okada, S.; Sugiyama, Y.; Kurai, K.; Iizuka, H.; Machida, A.; Miyakawa, Y.
J; Gen. Virol. 72, 2697-2704, 1991
R; Okamoto, H.; Okada, S.; Sugiyama, Y.; Kurai, K.; Iizuka, H.; Machida, A.; Miyakawa, Y.
J; Gen. Virol. 72, 2697-2704, 1991
A; Title: Nucleotide sequence of the genomic RNA of hepatitis C virus isolated from a hum A; Reference number: J01303
A; Molecule type: genomic RNA
A; Residues: 1-3033 < CKA>
A; Molecule type: genomic RNA
A; Residues: 1-3033 < CKA>
A; Cross-references: GB:D00944; NID:g221650; PIDN:BAA00792.1; PID:g221651
A; Estatus predicted < CPC>
A; Estatus predicted < CPC>
A; Estatus predicted < CPC>
F; 2-115/Product: capsid protein W #status predicted < MEE>
F; 2-115/Product: major envelope protein B *status predicted < NES>
F; J01-16-19/Product: nonstructural protein NSI *status predicted < NSI>
F; J01-16-19/Product: nonstructural protein RSI *status predicted < NSI>
F; J01-16-19/Product: nonstructural protein RSI *status predicted < NSI>
F; J01-16-19/Product: nonstructural protein RSI *status predicted < NSI>
F; J01-16-19/Product: nonstructural protein RSI *status predicted < NSI>
F; J01-16-19/Product: nonstructural protein RSI *status predicted < NSI>
F; J01-16-19/Product: nonstructural protein RSI *status predicted < NSI>
F; J01-16-19/Product: nonstructural protein RSI *status predicted < NSI>
F; J01-16-19/Product: nonstructural protein RSI *status predicted < NSI>
F; J01-16-19/Product: nonstructural protein RSI **
F; J01-16-19/Product: nonstructural RSI **
F; J01-16-19/Product: nonstructural RSI **
F; J01-16-19/Product: nonstructural RSI **
F; J01-16-19
                                                                                                                                                                                                                                A Note: the translation of the nucleotide sequence is not complete in this paper c) Superfamily: hepatitis C virus genome polyprotein
C; Superfamily: hepatitis C virus genome polyprotein
C; Superfamily: hepatitis C virus genome polyprotein
C; Superfamily: hepatitis C virus genome polyprotein
F; 2-115/Product: capsid protein C #status predicted <CPC>
F; 116-191/Product: capsid protein C #status predicted <CPC>
F; 116-191/Product: envelope protein E #status predicted <NED>
F; 1384-408/Region: hypervariable #status predicted vNSI>
F; 394-730/Product: nonstructural protein NS2 *status predicted <NS2>
F; 1008-161/Product: nonstructural protein NS2 *status predicted <NS2>
F; 1131-1138/Region: nucleotide-binding motif A (P-loop)
F; 1313-1318/Region: nucleotide-binding motif B
F; 1317-1320/Region: protein Protein NS2 *status predicted <NS3>
F; 1317-1320/Region: nucleotide-binding motif B
F; 1317-1320/Region: protein Protein NS2 *status predicted <NS3>
F; 1317-1320/Region: nucleotide-binding motif B
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      genome polyprotein - hepatitis C virus (isolate HC-J6)
N:Contains: capsid protein C; envelope protein M; hepacivirin (EC 3.4.21.98) (nonstructu
protein NS4a; nonstructural protein NS4b; nonstructural protein NS5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ä
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                113
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       114 CGSSDLYLVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAV 173
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           C:Species: hepatitis C virus
C:Date: 19-May-2000 *sequence_revision 19-May-2000 *text_change 17-Nov-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     54 AAQTFLATCINGVCWTVYHGAGTRTIASPKGPVIQMYTNVDKDLVGWPAPQGSRSLTPCT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        3 KKGSVVIVGRIN------LSGDTAYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVST
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      F.1617-1863/Product: nonstructural protein NS4a #stalus predicted <N4A>F.1864-2014/Product: nonstructural protein NS4b *stalus predicted <N4B>F.2015-3014/Product: nonstructural protein NS5 *stalus predicted <NS5>F.2210-2249/Region: interferon sensitivity determining *status predicted
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        F;1620-1866/Product: nonstructural protein NS4a *status predicted <N4A>F;1867-2017/Product: nonstructural protein NS4b *status predicted <N4B>
                                        A;Molecule type: mRNA
A;Residues: 1-3014 <CHA>
A;Cross-references: GB:Y13184
A;Experimental source: genotype 5a, which predominates in South Africa
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           DB 1; Length 3014;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Ouery Match 74.3%; Score 766.5; DB 1
Best Local Similarity 69.6%; Pred. No. 1.3e-62;
Matches 142; Conservative 25; Mismatches 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        174 CIRGVAKAVDFIPVESLETIMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Š
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
```

```
Micontains: capsid protein - hepatitis C virus (strain HC-J8)
Nicontains: capsid protein C; envelope protein M; hepacivirin (EC 3.4.21.98) (nonstructein NS4s; nonstructural protein NS4b; nonstructural protein NS5
C; Species: hepatitis C virus
C; Species: bapatitis C virus
C; Date: 31-Dec-1992 #sequence_revision 31-Dec-1992 #text_change 19-Jan-2001
C; Accession: A40250; PQ0397; PQ0559
R; Okamoto, H.; Kurai, K.; Okada, S.I.; Yamamoto, K.; Lizuka, H.; Tanaka, T.; Fukuda, Virology 188, 331-341, 1992
A; Title: Full-length sequence of a hepatitis C virus genome having poor homology to A; Reference number: A40250; MUID:92230232; PMID:1314459
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Across-references: GB:D10562, GB:D90518; NID:g221523; PIDN:BAA01418.1; PID:g221524
C:Superfamily: hepatitis C virus genome polyprotein
F:11-115/Product: capsid protein C *status predicted <CPC>
F:116-191/Product: envelope protein M *status predicted <NED>
F:130-733/Product: nonstructural protein NSI *status predicted <NSI>
F:134-1010/Product: nonstructural protein NSI *status predicted <NSI>
F:121-1215/Product: hepacivirin *status predicted <NSI>
F:121-1217/Region: nucleotide-binding motif A (P-loop)
F:1316-1321/Region: DEXH motif
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      A; Molecule type: genomic RNA
A; Residues: 1.3033 < CMA>
A; Residues: 1.3033 < CMA>
A; Cross-references: GB:D10988; GB:D01221; NID:g221608; PIDN:BAA01761.1; PID:g221609
A; Chan, S: W.; McCmish, F.; Holmes, E.C.; Dow, B.; Peutherer, J.F.; Follett, E.; Yap, J.; Gen. Virol. 73, 1131-1141, 1992
A; Title: Analysis of a new hepatitis C virus type and its phylogenetic relationship A; Reference number: PQ0393; MUID:92268871; PMID:136939
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                S.; Hijikata, M.; Shimot
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           F;1620-1866/Product: nonstructural protein NS4a #status predicted <N4A>
F;1867-2017/Product: nonstructural protein NS4b #status predicted <N4B>
F;2018-3033/Product: nonstructural protein NS5 #status predicted <NS5>
F;196,209,233,299,305,417,423,430,448,477,534,542,558,578,627,649,1091,1217,1259,203
                                                                                                                                                                                                                                                                                                                                                                                             ö
                                                                                                                                                                                                                                                              1034 TAYAQQTRGLLGTIVVSMTGRDKTEQAGEIQVLSTVTQSFLGTTISGVLWTVYHGAGNKT 1093
                                                                                          ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     79 IASPKGPVIQMYINVDKDLVGWPAPQGSRSLTPCTCGSSDLYLVTRHADVIPVRRRGDSR 138
                                                                                                                                                                        19 TAYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVSTAAQTFLATCINGVCWTVYHGAGTRT 78
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        139 GSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVCTRGVAKAVDFIPVESLETIMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Gaps
                                                                                               Gaps
                                                                                          ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     65.1%; Score 672; DB 1; Length 3033; 68.7%; Pred. No. 8.4e-54;
ive 24; Mismatches 32; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               A; Molecule type: genomic RNA
A; Residues: 2678-2754 <CHA>
A; Cross-references: DBJ:D10134
A; Experimental source: isolate E-b12
R; Kato, N.; Octsuyama, Y.; Ohkoshi, S.; Nakazawa, T.; Mori, Blochem. Biophys. Res. Commun. 181, 279-285, 1991
A; Title: Distribution of plural HCV types in Japan.
A; Reference number: PQ0554; MUID:92068204; PMID:1720309
A; Accession: PQ0559
65.3%; Score 674; DB 1; 68.2%; Pred. No. 5.5e-54;
                                                                                 26; Mismatches
```

qq ολ a ö g 12

UΣ

```
anglo-associated migratory cell protein - human c.Species: Homo sapiens (man) c.Species: Homo sapiens (man) c.Species: Homo sapiens (man) c.Species: Homo sapiens (man) c.Species: No. Sep. 1996 *text_change 21-Jul-2000 c.Specession: I39383 R.Beckner, M.E.; Krutzsch, H.C.; Stracke, M.L.; Williams, S.T.; Gallardo, J.A.; Liott Cancer Res. 55, 2140-2149, 1995 A.Title: Identification of a new immunoglobulin superfamily protein expressed in bloo A.Reference number: I39383; MUID:95262124; PMID:7743515
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  A;Residues: 1-495 <COL>
A;Cross-references: GB:AE001199; GB:AE000520; NID:g3322402; PIDN:AAC65137.1; PID:g332 A;Experimental source: strain Nichols C;Genetics:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Pypothetical protein TP0136 - syphilis spirochete
Cippecies: Treponema palidum subsp. palidum (syphilis spirochete)
Cipate: 24-7ul-1998 #sequence_revision 24-7ul-1998 #text_change 05-Nov-1999
Cipate: 24-7ul-1998 #sequence_revision 24-7ul-1998 #text_change 05-Nov-1999
Cipate: 24-7ul-1998 #sequence_revision 24-7ul-1998 #text_change 05-Nov-1999
Cipate: 24-7ul-1998 #sequence 07-7ul-1998 #text_change 05-Nov-1999
Cipate: 24-7ul-1998 #sequence 07-7ul-1998 #text_change 05-Nov-1999
Cipate: 24-7ul-1998 #sequence of Treponema pallidum, the syphilis spirochete.
A:Reference number: A71250: MUID:98332770; PMID:9665876
                                                                                 1066 KCQPQGVWVI-----RND--GALCHGTLGRTVELDLPAELCDFRGSSGSPILCD 1112
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                64 AGSKLYATNGRLWEKELNGTGSWQKVSSSSVPTDSDK----KVMSIATDGNTFVLACVP 118
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       119 GTGVYKHCVNGAGSSSTGTTASPSTETCSQHAT····LVGGTSKPFWLVPGGTGNNGNCG 174
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          114 C-----GSSDLFLVTRHADVIP------VRRRGDSRGSLLSPRPISYLK------ 151
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          175 CGGGGGGSSSSSSCIHIWLVPGGTGNNGNCGCGGGGGSSSSSSSCIHIKVENTDEQFL 234
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               A;Accession: B71360
A;Status: preliminary; nucleic acid sequence not shown; translation not shown A;Molecule type: DNA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    64 -NGVCWTVYHGAG---TRTIASPKGPVIQMYTNVDKDLVG-----WPAPQGSRSLTPCT
                                          53 TAAQIFLATCINGVCWTVYHGAGTRIIASPKGPVIQMYINVDKDLVGWPAPQGSRSLTPC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    16 SGDTAYA------QQTRGEEGCQETSQ----TGRDKNQVEGEVQIVSTAAQTFLATCI-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               A; Cross-references: GB: M95627; NID: 9870802; PIDN: AAA68889.1; PID: 9870803
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              8.3%; Score 85.5; DB 2; Length 495; llarity 23.4%; Pred. No. 3.7; Conservative 17; Mismatches 76; Indels 7
                                                                                                                                                                                                                                                                                                                162 AGHAVGIFRAAVCTRG-----VAKAVDFIPVESLETIMRSP 197
                                                                                                                                                                                                                                                                                                                                                          -----GSSGGPLLCPAGHAVG 167
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                DMGEGYVVTTKHLYTKNGSSSAGPAQCPGGGGG 268
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-452 <RES>
                                                                                                                                                                             113 TCGSSDLYLVTRHADVÍPVRRRGDSRGSLLS---
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Best Local Similarity
Matches 50; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                A; Accession: I39383
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  C;Genetics:
A;Gene: TP0136
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         152
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                235
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Genetics:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      RESULT 14
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      a
                                                                                                      a
                                                                                                                                                                                                                                                a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                   ð
                                                                                                                                                                                                                                                                                                                                                                                셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        polyprotein - marmoset hepatitis GB virus A
C;Species: 13-01-1999 *sequence_revision 23-01-1999 *text_change 17-Nov-2000
C;Accession: T08839
B;Erker, J.C.; Desai, S.M.; Leary, T.P.; Chalmers, M.L.; Montes, C.C.; Mushahwar, I.K.
J. Gen. Virol. 79, 41-45, 1998
A;Title: Genomic analysis of two GB virus A variants isolated from captive monkeys.
A;Reference number: 216486; MUID:98120818; PMID:9460920
A;Status: translated from GB/EMBL/DDBJ
A;Molecule type: genomic RNA
A;Residues: 1-2970 cRRKx
A;Residues: 1-2970 cRRKx
A;Residues: 1-2970 cRRKx
C;Superfamily: hepatitis C virus genome Polyprotein
C;Reywords: polyprotein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           polyprotein - douroucouli hepatitis GB virus A
C:Species: douroucouli hepatitis GB virus A
C:Species: douroucouli hepatitis GB virus A
C:Date: 20.5ep-1999 #sequence_revision 20.5ep-1999 #text_change 17-Nov-2000
C:Date: 20.5ep-1999 #sequence_revision 20.5ep-1999 #text_change 17-Nov-2000
C:Accession: T08841
A:Erker, J.C.; Desai, S.M.; Leary, T.P.; Chalmers, M.L.; Montes, C.C.; Mushahwar, I.K.
J. Gen. Virol. 79, 41-45, 1998
A;Title: Genomic analysis of two GB virus A variants isolated from captive monkeys.
A;Reference number: 216486; MUID:98120818; PMID:9460920
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          'n
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          995 KTSMLGRDEREHEGSIVVLGTSTTRSMGTCVNGVMYTTFHGSNARTLAGPVGPVNCRWWS 1054
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1055 PSDDVAVYPLPSGASCLEPCKCGTQSVWCIRN--DGALCHGRLSKLVELDLPTEISDFRG 1112
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             A:Molecule type: mRNA
A:Residues: 1-3005 <ERK>
A:Cross-references: EMBL:AF023425; NID:g2828599; PIDN:AAC40502.1; PID:g2828600
C;Superfamily: hepatitis C virus genome polyprotein
C;Keywords: polyprotein
                                                                                                                                      79 IASPKGPVIQMYINVDKDLVGWPAPQGSRSLIPCTCGSSDLYLVTRHADVIPVRRRGDSR 138
                                                                                                                                                                                                                                                                                                                   93 VDKDLVGWPAPQGSRSLTPCTCGSSDLYLVTRHADVIPVRRRGDSRGSLLSPRPISYLKG 152
   TAYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVSTAAQTFLATCINGVCWTVYHGAGTRT 78
                                                                                                                                                                                                                                                                            GSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVCTRGVAKAVDFIPVESLETTMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          33 ETSQIGRDKNQVEGEVQIVSIAAQIFLATCINGVCWIVYHGAGIRIIASPKGPVIQMYIN 92
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            3 KKGSVVIVGRIN-----LSGDTAYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVS 52
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     10;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Length 3005;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Query Match 24.9%; Score 257; DB 2; Length 300
Best Local Similarity 34.1%; Pred. No. 3.4e-15;
Matches 56; Conservative 29; Mismatches 69; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SSGGPLLCPAGHAVGIFRAAVCTRGV-----AKAVDFIPVES 189
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 A:Status: translated from GB/EMBL/DDBJ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Query Match 24.3%
Best Local Similarity 27.8%
Matches 62; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                153
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1113
19
                                                                                                                                                                                                                                                                            139
```

à 셤 ð g ò g

윱 ð

Search completed: August 30, 2003, 19:20:27 Job time : 17.2134 secs

```
GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.
```

OM protein - protein search, using sw model

August 30, 2003, 18:01:52 ; Search time 9.75674 Seconds (without alignments) 949.524 Million cell updates/sec Run on:

US-09-965-594-14 1032 1 MKKKGSVVIVGRINLSGDTA.....VAKAVDFIPVESLETIMRSP 197 Title: Perfect score: Sequence:

BLOSGM62 Gapop 10.0 , Gapext 0.5 Scoring table:

127863 seqs, 47026705 residues Searched:

127863 Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0 Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0% Maximum Match 100% Listing first 45 summaries

SwissProt\_41:\* Database : Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## STIMMARTES

		di			SUMMARIES		
Result No.	Score	Query Match	Query Match Length	DB	ID	Description	
	1 100	0.7		-			
4	201.3		2011	-	PODG_RCV1		
7	886.5	85.9	3011	-	POLG_HCVH	c	
٣	854.5	2	3010	-	POLG_HCVTW	c	
4	850.5	3	3010	-	POLG_HCVBK	h genome	
2	850.5	4	3010	-	POLG_HCVJA	h genome	
9	844.5	ä	3010	-	POLG_HCVJT	h genome	
7	674	65.3	3033	٦	POLG_HCVJ6		
co	672	5.	3033	-	POLG_HCVJ8	h genome	
6	85.5		485	-	Y136_TREPA	treponema	
<b>1</b> 0	85		321	Н	HHOA_ARATH		
11	82		452	~	AAMP_HUMAN		
12	80.5		437	-	DEGLARATH	arabi	
13	78.5		209	-	PAAD_PSEAE		
14	77.5		2663	-	CENE_HUMAN	002224 homo sapien	
15	16		911	٦	TB11_NEIMB	neiss	
	75.5	7.3	786	7	SN1L_HUMAN	P57059 homo sapien	
	75		603		ENV_RSVP	rons	
18	7		1705	7	PTPO_MOUSE		
19	4		415	~	ZP3_RABIT	P48833 oryctolagus	
20	74.5		116	7	HYPF_A20VI		
21	~	•	1165	-	POL_GALV	gib	
22	73.5	٠	263	Н	GRAK_MOUSE	_	
23	ë.	7.1	199	~		dauc	
24	73	٠	253	-	CAC3_BOVIN	P05805 bos taurus	
25	73	٠	259	Н	IBP1_HUMAN	homo	
56	72.5	•	257	7	GRAM_HUMAN	рошо	
27	72.5		206	-	TRFE_HORSE	ednns	
28	72		629	-	VST2_HEVME	Q03500 hepatitis e	
29	~	•	1527	-	CA1H_MOUSE	_	
30			248	-		P11033 mus musculu	
31	71.5	6.9	248	-	TRY1_CHICK	Q90627 gallus gall	
32		6.9	408	-			
	71		336	-	UL16_EBV	epstein	

001996 neisseria g P18614 rattus norv P28614 rattus norv 909550 caenorhabdi 909550 caenorhabdi 9067531 aquifex aeo 908359 feline leuk P29326 hepatitis e P33426 hepatitis e P19532 homo sapien	05/5 pseudomonas
4 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0	P.20
CN4A_RAT TBPL_NEIGO ITAL_RAT YQU3_CAEEL CTEL_HUMAN FLIT_AQUAE ENV_FLVGL VST2_HEVBU VST2_HEVBU TFE3_HUMAN	IKFU_PSEPU
напананана п	٦ -
88 11180 11280 1280 1280 1280 1280 1280	0.4. V
00000000000000000000000000000000000000	0
71 71 70 70 70 70 70 70 70 70 70 70 70	2
<pre></pre>	n F

## ALIGNMENTS

```
of unwinding.";
Structure 6:89-100(1998)
                                                                                                                                                                                                                                                                                                                                                                          STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ACTIVATION OF NS3.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     NCBI_TaxID-11108;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Prince A.M.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Hepacivirus
                                                                                                                                                                                                                                                                                                                                                                          POLG_HCVH
                                                            ò
                                                                                                   g
                                                                                                                                               ò
                                                                                                                                                                                 g
                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CELLULAR ANINOPEPTIDASE.
CARSID PROTEIN C (POTENTIAL).
MATRIX PROTEIN (POTENTIAL).
MAJOR ENVELOPE PROTEIN E (POTENTIAL).
NONSTRUCTURAL PROTEIN NS1/F2 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS2 (POTENTIAL).
PROTEASE/HELICASE NS3 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4A (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4A (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4A (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Polyprotein; Glycoprotein; Transferase; RNA-directed RNA polymerase;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Core protein; Coat protein; Envelope protein; Helicase; ATP-binding;
Transmembrane; Nonstructural protein; Hydrolase; Serine protease;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               RNA-DIRECTED RNA POLYMERASE (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       REMOVED FROM CAPSID PROTEIN C BY THE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CHARGE RELAY SYSTEM (BY SIMILARITY)
CHARGE RELAY SYSTEM (BY SIMILARITY)
CHARGE RELAY SYSTEM (BY SIMILARITY)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              87.4%; Score 901.5; DB 1; Length 3011;
.larity 85.8%; Pred. No. 5.5e-77;
Conservative 9; Mismatches 11; Indels 9;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  (POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       (POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         65F8C9447FCE5AF9 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              N-LINKED (GLCNAC.
N-LINKED (GLCNAC.
N-LINKED (GLCNAC.
N-LINKED (GLCNAC.
N-LINKED (GLCNAC.
N-LINKED (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (GLCNAC.
(GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (GLCNAC.
(GLCNAC.
(GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  N-LINKED (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ATP (POTENTIAL)
DECH BOX.
                                ...v_NS1.
...v; IPR004109; HCV_NS2.
InterPro; IPR004109; HCV_NS3.
InterPro; IPR001490; HCV_NS4.
DR InterPro; IPR001286; HCV_NS4b.
DR InterPro; IPR001266; HCV_NS5a.
DR InterPro; IPR00156; HCV_RGAP.
DR InterPro; IPR001059; RNA_Dol_DS_PS.
'R InterPro; IPR007094; RNA_Dol_DS_PS.
'R InterPro; IRR007094; RNA_Dol_DS_PS.
'R Pfam; PF01543; HCV_Capsid; 1.
Pfam; PF01542; HCV_Core; 1.
Pfam; PF01542; HCV_Core; 1.
Pfam; PF015567.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          2240 2240 N-L.
2364 2364 N-L.
2789 2789 N-L.
3011 AA: 327197 MW;
                                                                                                                                                                                                                                                                                                                                                  Pfam; PF02907; HCV_NS3; 1.
Pfam; PF01006; HCV_NS4s; 1.
Pfam; PF01001; HCV_NS4s; 1.
Pfam; PF01001; HCV_NS5s; 1.
Pfam; PF00271; helicase_C; 1.
Pfam; PF00998; viral_RdRP; 1.
ProDom; PD186062; HCV_NS1; 1.
                                                                                                                                                                                                                                                                                   Pfam; PFULSC.,
Pfam; PF01560; HCV_NSI,
Pfam; PF01539; HCV_NS2; 1.
Pfam; PF01006; HCV_NS3; 1.
Pfam; PF01006; HCV_NS4s; 1.
                       HCV_env.
HCV_NS1.
                     IPR002519;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      175; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    3D-structure
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CARBOHYD
CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ACT_SITE
ACT_SITE
ACT_SITE
NP_BIND
SITE
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CARBOHYD
                          InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SECUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         INIT_MET
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Best Local
Matches 17
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CHAIN
CHAIN
CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CHAIN
SOLITIES SELECTED SEL
```

53

Gaps

6 .

3 KKGSVVIVGRIN-----LSGDTAYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVST

ò

```
54 AAQTFLATCINGVCWTVYHGAGTRTIASPKGPVIQMYTNVDKDLVGWPAPQGSRSLTPCT 113
                                                                                                                                                                                                                                                                                                                                                                                                                                    01-AUG-1992 (Rel. 23, Created)
01-AUG-1992 (Rel. 23, Last sequence update)
01-AUG-1992 (Rel. 23, Last sequence update)
01-AUG-1992 (Rel. 24, Last annotation update)
6enome polyprotein (Contains: Capsid protein C (Core protein) (P22);
6enome polyprotein E1 (GP32) (GP35); Envelope glycoprotein E2
(GP68) (GP70) (NS1); Protein P7; Nonstructural protein NS2 (P21)
(EC 3.4.99-); Protease/helicase NS3 (P70) (Hepacivirin)
(EC 3.4.198); Nonstructural protein NS48 (P4); Nonstructural protein
NS48 (P27); Nonstructural protein NS5A (P56); Nonstructural protein
NS5B (P66) (P70) (RMA-directed RNA polymerase) (EC 2.7.7.48)].
Hepatitis C virus (isolate H) (HCV).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        X-RAY CRYSTALLOGRAPHY (2.1 ANGSTROMS) OF 1207-1657.
MEDLINE-9733132; PubMed-9187654;
Yao N., Hesson T., Cable M., Hong Z., Kwong A.D., Le H.V., Weber P.C.;
"Structure of the hepatitis C virus RNA helicase domain.";
Nat. Struct. Biol. 4:463-467(1997).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS) OF 1192-1657.
MEDLINE-98154321; PubMed-9493270;
Rim J.L., Morgenstern R.A., Griffith J.P., Dwyer M.D., Thomson J.A.,
Murcko M.A., Lin C., Caron P.R.,
"Hepatitis C virus NS3 RNA helicase domain with a bound
oligonucleotide: the crystal structure provides insights into the mode
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      -I- FUNCTION: PROTEASE NS2 IS RESPONSIBLE FOR THE CLEAVAGE OF NS2-NS3-I-FUNCTION: PROTEASE NS3 IS RESPONSIBLE FOR THE CLEAVAGE OF NS3-NS3-NS4A, NS4A-NS4B, NS4B-NS5A AND NS5A NS5A NS5B.
-I-FUNCTION: NS4A FORMS A COMPLEX WITH NS3 AND IS ESSENTIAL FOR THE
                                                                                    CATALYTIC ACTIVITY: Hydrolysis of four peptide bonds in the viral precursor polyprotein, commonly with Asp or Glu in the P6 position, Cys or Thr in P1 and Ser or Ala in P1.

CATALYTIC ACTIVITY: N nucleoside triphosphate - N diphosphate +
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           (RNA)(N).
SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A SUBUNIT: THE VIRIOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS: EL AND E2. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MRNA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            -1- FUNCTION: NSSA SEEMS TO HAVE A TRANSCRIPTIONAL ACTIVATORY ROLE.
-1- FUNCTION: NSSB IS A RNA-DEPENDENT RNA POLYMERASE THAT PLAYS AN ESSENTIAL ROLE IN THE VIRUS REPLICATION.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ပ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  "Genomic structure of the human prototype strain H of hepatitis virus: comparison with American and Japanese isolates."; Proc. Natl. Acad. Sci. U.S.A. 88:10292-10296(1991).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SEQUENCE FROM N.A.
MEDLINE-92052256; PubMed-1658800;
Inchauspe G., Zebedee S., Lee D.H.H., Sugitani M., Nasoff M.,
                                                                                                                                                                                                                                                                                                                                                                                                   PRT; 3011 AA
                                                                                                                                                                                                                                                                 174 CTRGVAKAVDFIPVESLETTMRSP 197
```

```
(POTENTIAL)
                                                                                                                         (POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Length 3011;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                327142 MW; 772CBB29CCD94753 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Score 886.5; DB 1;
Pred. No. 1.5e-75;
                                             (GLCNAC.
(GLCNAC.
(GLCNAC.
(GLCNAC.
(GLCNAC.
(GLCNAC.
              (GLCNAC.
(GLCNAC.
(GLCNAC.
   N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    85.9%;
83.8%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                               1417
1420
1436
1439
1453
1463
1478
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1597
1598
1611
1618
1623
1627
1627
1636
                                                                                                                                                                                                                                                                                                                                                                                                                                           1409
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ¥.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Query Match
Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                               1414
1419
1432
1438
1450
1456
1471
1480
1489
              STRAND
STRAND
STRAND
HELIX
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    STRAND
STRAND
STRAND
STRAND
STRAND
HELIX
TURN
STRAND
STRAND
STRAND
HELIX
HELIX
HELIX
HELIX
HELIX
HELIX
HELIX
HELIX
HELIX
                                                                                                                                                                                                                                                                                                                                                                                             STRAND
HELIX
STRAND
HELIX
TURN
STRAND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        HELIX
FORN
HELIX
FORN
    This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation-the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Polyprotein; Glycoprotein; Transferase; RNA-directed RNA polymerase; Core protein; Coat protein; Envelope protein; Helicase; ATP-binding; Transmembrane; Nonstructural protein; Hydrolase; Serine protease; 10.5tructure. 1 REMOVED FROM CAPSID PROTEIN C BY THE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SIMILARITY).
SIMILARITY).
SIMILARITY).
PTM: THE STRUCTURAL PROTEINS C, E1 AND E2 ARE PRODUCED BY PROTEOLYTIC PROCESSING BY THE HOST SIGNAL PEPTIDASES. SIMILARITY: THE NS2 PROTEASE BELONGS TO PEPTIDASE FAMILY U39. SIMILARITY: THE NS3 PROTEASE BELONGS TO PEPTIDASE FAMILY S29.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       N-LINKED (GLCNAC. .) (POTENTIAL).
N-LINKED (GLCNAC. .) (POTENTIAL).
N-LINKED (GLCNAC. .) (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             NS4A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             PROTEIN P7.
NONSTRUCTURAL PROTEIN NS2.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ENVELOPE GLYCOPROTEIN E1. ENVELOPE GLYCOPROTEIN E2.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CELLULAR AMINOPEPTIDASE. CAPSID PROTEIN C.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CHARGE RELAY SYSTEM (BY
CHARGE RELAY SYSTEM (BY
CHARGE RELAY SYSTEM (BY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  PROTEASE/HELICASE NS3.
NONSTRUCTURAL PROTEIN NS
NONSTRUCTURAL PROTEIN NS
NONSTRUCTURAL PROTEIN NS
NONSTRUCTURAL PROTEIN NS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ATP (POTENTIAL).
DECH BOX.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        POTENTIAL.
                                                                                                                                                                                                                 TRANSFAC: 193:01;
InterPro: IPR001410;
InterPro: IPR00155;
InterPro: IPR001521; HCV_capsid.
InterPro: IPR002521; HCV_capsid.
InterPro: IPR002519; HCV_core.
InterPro: IPR002519; HCV_core.
InterPro: IPR002519; HCV_NS1.
InterPro: IPR00199; HCV_NS2.
InterPro: IPR001099; HCV_NS3.
InterPro: IPR001409; HCV_NS4.
InterPro: IPR001409; HCV_NS4.
InterPro: IPR001409; HCV_NS4.
InterPro: IPR001056; HCV_NS3.
InterPro: IPR001059; RNA_pol_DS_PSA.
InterPro: IPR001059; RNA_pol_DS_PSA.
InterPro: IPR001059; HCV_core; I.
Pfam; PF01539; HCV_core; I.
Pfam; PF01539; HCV_NS3; I.
Pfam; PF01506; HCV_NS3; I.
Pfam; PF01001; HCV_NS4; I.
Pfam; PF00019; HCV_NS5: I.
Pfam; PF00019; HCV_NS5: I.
Pfam; PF00099; VITAL_RRP; I.
PRODOM; PD186062; HCV_NS1; I.
                                                                                                                                           EMBL; M67463; AAA45534.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            191
383
746
809
809
809
1057
11711
1711
1711
1083
1083
1108
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
                                                                                                                                                    PIR; A36814; GNWVCH.
PDB; 1HEI; 25-NOV-98.
PDB; 1AIV; 16-FEB-99.
PDB; 1AIR; 17-JUN-98.
MEROPS; 229.001;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CHAIN
CHAIN
CHAIN
CHAIN
CHAIN
CHAIN
CHAIN
CHAIN
ACT_SITE
ACT_SITE
ACT_SITE
ACT_SITE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CARBOHYD
CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CHAIN
```

```
InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                      CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                 CARBOHYD
                                                                                                                                                                                                                                                                                                                              CHAIN
                                                                                                                                                                                                                                                                                                                                                        CHAIN
                                                                                                                                                                                                                                                                                                   CHAIN
                                                                                                                                                                                                                                                                                   CHAIN
                                                                                                                                                                                                                                                                                                                       CHAIN
                                                                                                                                                                                                                                                                                                                                                 CHAIN
  This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entitles requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                             01-APR-1993 (Rel. 25, Created)
01-APR-1993 (Rel. 25, Last sequence update)
01-APR-1993 (Rel. 25, Last sequence update)
01-PRP-1993 (Rel. 42, Last annotation update)
Genome polyprotein (Contains notation and protein (P22);
Envelope glycoprotein El (GP32) (GP35); Envelope glycoprotein E2
(GP68) (GP70) (NS1); Protein P1 Nonstructural protein NS2 (P21)
(EC 3 4.21.98); Nonstructural protein NS48 (P4); Nonstructural protein NS48 (P27); Nonstructural protein NS58 (P56); Nonstructural protein NS58 (P66) (P70) (RNA-directed RNA polymerase) (EC 2.7.7.48)].
Hepatitis C virus (isolate Taiwan) (HCV).
                                                                                                                                                                                                                                                                                                                                             54 AAQTFLATCINGVCWTVYHGAGTRTIASPKGPVIQMYTNVDKDLVGWPAPQGSRSLTPCT
                                                                                          CGSSDLYLVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAV
                   3 KKGSVVIVGRIN-----LSGDIAYAQQIRGEEGCQETSQIGRDKNQVEGEVQIVST
                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SUBMIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A LIPOPROTEIN RAVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS: PROTEIN M AND GLYCOPROTFIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C. AND MRNA.
.
ف
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY 529.
14; Indels
                                                                                                                                                                                           PRT; 3010 AA
Mismatches
                                                                                                                           EMBL; M84754; -; NOT_ANNOTATED_CDS
10;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             InterPro; IPR001410; DEAD.
Matches 171; Conservative
                                                                                                                                                                                            STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          PDB; 1N64; 25-FEB-03
PDB; 1NS3; 08-APR-98
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                PIR; A40244; GNWVTW
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           MEROPS; S29.001;
MEROPS; U39.001;
                                                                                                                                                                                                                                                                                                                             NCBI_TaxID=31645;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                 {RNA}(N)
                                                                                                                                                                                           POLG_HCVTW
P29846;
                                                                                                          1125
                                                                                         114
                                                                                                                                                                                   POLG_HCVTW
                                                                                                                                                                         RESULT 3
                                                                                                                                                                                            a
                                                                                                                            ö
                                                       õ
                                                                        a
                                                                                     ò
                                   g
```

```
ij
              MATRIX PROTEIN (POTENTIAL).
MAJOR ENVELOPE PROTEIN E (POTENTIAL).
NONSTRUCTURAL PROTEIN NSI/E2 (POTENTIAL).
PROTEASE/HELICASE NS3 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS44 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS48 (POTENTIAL).
RNA-DIRECTED RNA POLYMERASE (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Gaps
                                                                                                                                                                                                                                                                                                                                                                                        SIMILARITY).
SIMILARITY).
SIMILARITY).
                                                                                                                                                                                                                                                                                                                                                                                                                                               (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ä
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Length 3010;
                                                                                                                                                                                                                                                                        REMOVED FROM CAPSID PROTEIN C BY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              327047 MW; AAD267D55CDFE215 CRC64;
                                                                                                                                                                                                                                                                                  CELLULAR AMINOPEPTIDASE. CORE PROTEIN (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                    N-LINED GLCNAC...)
                                                                                                                                                                                                                                                                                                                                                                                        CHARGE RELAY SYSTEM (BY
CHARGE RELAY SYSTEM (BY
CHARGE RELAY SYSTEM (BY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  DB 1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     N-LINKED (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (GLCNAC. . . (GLCNAC. . . (GLCNAC. . .
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Score 854.5; DB ]
Pred. No. 1.6e-72;
                                                                                                                                                                                                                                                                                                                                                                                                                     ATP (POTENTIAL).
DECH BOX.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Mismatches
                                                                                                                                                                                                                                                                                                                                                                                 POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      18;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  82.8%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Conservative
                                                                                                                                                                                                                                                                                           IPR002521;
IPR002519;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ¥.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Similarity
                                                                                                                                                                                                                                                               3D-structure
INIT_MET
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Query Match
Best Local Simi
Matches 161;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CARBOHYD
CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CARBOHYD
CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CARBOHYD
CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CARBOHYD
CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                               TRANSMEM
ACT_SITE
ACT_SITE
ACT_SITE
ACT_SITE
NP_BIND
                                                                                                                                                                                                                                                                                                                                                                                                                                                 CARBOHYD
```

us-09-965-594-14.rsp

```
1005 RRGREILLGPADSLEGRGWRLLAPITAYAQQTRGLFGCIITSLTGRDRNQVEGEVQVVST 1064
                                                             54 AAQTFLATCINGVCWTVYHGAGTRTIASPKGPVIQMYTNVDKDLVGWPAPQGSRSLTPCT 113
                                                                                              114 CGSSDLYLVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAV 173
3 KKGSVVIVGRIN------LSGDTAYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVST 53
                                                                                                                                                                                                                                                                   01-AUG-1992 (Rel. 23, Last Sequence update)
15-SEP-2003 (Rel. 42, Last annotation update)
66nome polyprotein [Contains: Capsid protein C (Core protein) (P22);
Envelope glycoprotein [Contains: Capsid protein C (Core protein E2 (GP68) (GP70) (NS1); Protein P7; Nonstructural protein NS2 (P21) (EC 3.4.22.-); Proteases/helicase NS3 (P70) (Hepaclytin)
(EC 3.4.21.98); Nonstructural protein NS4A (P4); Nonstructural protein NS5A (P56); Nonstructural protein NS5A (P56); Nonstructural protein NS5A (P56); Nonstructural protein NS5A (P56); Nonstructural protein NS5A (P56) (P70) (RAA-directed RNA polymerase) (EC 2.7.7.48)].
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Takamizawa A., Mori C., Fuke I., Manabe S., Murakami S., Fujita J., Onishi E., Andoh T., Yoshida I., Okayama H.:
"Structure and organization of the hepatitis C virus genome isolated
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      MEDLINE-97015088; PubMed-8861916;
Love R.A., Parget H.E., Wickersham J.A., Hostomsky Z., Habuka N.,
Moomaw E.W., Adachi T., Hostomska Z.;
"The crystal structure of hepatitis C virus NS3 proteinase reveals a
                                                                                                                                                                                                                                                                                                                                                                              Hepatitis C virus (isolate BK) (HCV).
Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS) OF 1027-1210 AND 1678-1691.
MEDLINE-98227846; PubMed-9568891;
Yan Y., Li Y., Munshi S., Sardana V., Cole J.L., Sardana M.,
Stafnkuehler C., Tomei L., de Francesco R., Kuo L.C., Chen Z.;
"Complex of NS3 protease and NS4A peptide of BK strain hepatitis C virus: a 2.2-A resolution structure in a hexagonal crystal form.";
Protein Sci. 7:837-847(1998).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   WEDLINE-96235224; PubMed-8647104;
BOTOWSKI P., Helland M., Oehlmann K., Becker B., Korneteky L.;
"Non-structural protein 3 of hepatitis C virus inhibits
phosphoxylation mediated by cAMP-dependent protein kinase.";
Eur. J. Biochem. 237:611-618(1996).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       trypsin-like fold and a structural zinc binding site."; Cell 87:331-342(1996).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            X-RAY CRYSTALLOGRAPHY (2.4 ANGSTROMS) OF 1027-1215.
                                                                                                                                                           174 CIRGVAKAVDFIPVESLETIMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                          SEQUENCE FROM N.A.
MEDLINE-91140698; Pubmed-1847440;
                                                                                                                                                                                                                                                          01-AUG-1992 (Rel. 23, Created)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       from human carriers.";
J. Virol. 65:1105-1113(1991).
                                                                                                                                                                                                                                   STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SEQUENCE OF 1487-1500.
                                                                                                                                                                                                                                                                                                                                                                                                                   NCBI_TaxID=11105;
                                                                                                                                                                                                                                                                                                                                                                                                        Hepacivirus
                                                                                                                                                                                                                                  POLG_HCVBK
                                                                                                                                                                                                                      POLG_HCVBK
                                                                                                                                                                                                           RESULT
                                                                                                                                                                                                                                   a
                                                                      g
                                                                                              ò
                                                                                                                      a
                                                                                                                                               ò
                                                                                                                                                                    셤
                                                ò
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss institute of Bioinformatics and the EMED outstation the European Bioinformatics institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CAPSID PROTEIN C (POTENTIAL).
MAJOR ENVELOPE PROTEIN (POTENTIAL).
NONSTRUCTURAL PROTEIN E (POTENTIAL).
NONSTRUCTURAL PROTEIN NSI/E2 (POTENTIAL).
PROTEIN NSI/E2 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS2 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4A (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4A (POTENTIAL).
  BY A
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                RNA-DIRECTED RNA POLYMERASE (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Polyprotein, Glycoprotein, Transferase, RNA-directed RNA polymerase, Core protein, Coat protein; Envelope protein; Helicase; ATP-binding; Transmembrane; Nonstructural protein; Hydrolase; Serine protease;
SUBDUIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID CUVERED BI
LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF
PROTEIN C AND MRNA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                REMOVED FROM CAPSID PROTEIN C BY THE
                                                                                   SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY S29.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CELLULAR AMINOPEPTIDASE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CHARGE RELAY SYSTEM.
CHARGE RELAY SYSTEM.
CHARGE RELAY SYSTEM.
                                                                                                                                                                                                                                                                             ATP (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     InterPro; IPR00145; HCV_NS4.
InterPro; IPR00146; HCV_NS4.
InterPro; IPR001466; HCV_NS4.
InterPro; IPR001266; HCV_RGAP.
InterPro; IPR001266; HCV_RGAP.
InterPro; IPR007094; RNA_POL_DS_PS.
InterPro; IPR007094; RNA_POL_DS_PS.
InterPro; IPR001994; RNA_POL_PSVII.
Pfam; PP01543; HCV_cappid; 1.
Pfam; PP01543; HCV_cappid; 1.
Pfam; PP01559; HCV_NS1; 1.
Pfam; PP01580; HCV_NS2; 1.
Pfam; PP01506; HCV_NS3; 1.
Pfam; PP01006; HCV_NS3; 1.
Pfam; PP01006; HCV_NS4; 1.
Pfam; PP01006; HCV_NS4; 1.
Pfam; PP0100998; Viral_RGAP; 1.
Pfam; PP0100998; Viral_RGAP; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  HCV_capsid.
HCV_core.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   IPR002519; HCV_env.
IPR002531; HCV_NSI.
IPR002518; HCV_NS2.
IPR004109; HCV_NS3.
                                                                                                                                                                                                                                                                                                 EMBL; M58335; AAA72945.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  369
1083
1107
1165
1237
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          191
383
729
1006
1615
1615
2013
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SMART; SM00487; DEXDC;
                                                                                                                                                                                                                                                                                                                                      1A1Q; 25-MAR-98
1JXP; 14-JAN-98
1DX3; 08-APR-98.
1C2P; 15-NOV-00
1CSJ; 08-NOV-99.
1GX5; 09-APR-02.
1GX6; 10-APR-02.
1GX6; 26-JUN-00.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            InterPro; IPR00141Q;
                                                                                                                                                                                                                                                                                                                     GNWVTC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1007
1616
1863
2014
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                $29.00
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        u39.00
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             3D-structure.
INIT_MET
                                                                                                                                                                                                                                                                                                                       A38465;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  InterPro;
InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                interPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  TRANSMEM
ACT_SITE
ACT_SITE
ACT_SITE
ACT_SITE
NP_BIND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    MEROPS;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      MEROPS;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CHAIN
CHAIN
CHAIN
CHAIN
CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CHAIN
                                                                                                                                                                                                                                                                                                                                                               PDB;
PDB;
PDB;
PDB;
PDB;
  --
```

1185 CTRGVAKAVDFVPVESMETTMRSP 1208

```
1064
                                                                                                                                                                                                                                                                                                                                                                                                              114 CGSSDLYLVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAV 173
                                                                                                                                                                                                                                                                                                                                                                                       54 AAQIFLATCINGVCWTVYHGAGIRTIASPKGPVIQMYTNVDKDLVGWPAPQGSRSLIPCT
                                                                                                                                                                                                                                                                                                                                                                               3 KKGSVVIVGRIN-----LSCDTAYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVST
                                                                                                                                                                                                                                                                                                                                                                    Gaps
                                                 (POTENTIAL).
(POTENTIAL).
(POTENTIAL).
                                                                    (POTENTIAL)
(POTENTIAL)
(POTENTIAL)
(POTENTIAL).
(POTENTIAL).
                               (POTENTIAL).
(POTENTIAL).
(POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                  .5
                          (POTENTIAL)
                                                                                                                              (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                       Length 3010;
                                                                                                                                                                                                                                                                                                                                                                   Indels
                                                                                                                                                                                                                                                                                                                                         327189 MW; F8422D5ECCFDFD9C CRC64;
                                                                                                                                                                                                                                                                                                                                                     82.4%; Score 850.5; DB 1; 77.5%; Pred. No. 3.8e-72; live 21; Mismatches 16;
                                                         (GLCNAC.
                  (GLCNAC.
                                     GLCNAC.
                                                  GLCNAC.
                                                                                       GLCNAC.
                                                                                               (GLCNAC.
                                                                            (GLCNAC.
                                                                                                           (GLCNAC.
                                (GLCNAC.
                                                                                                    GICNAC.
                                                                    (GLCNAC.
                                                                                 (GLCNAC.
      N-LINKED
                                                                                                   N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                           CTRGVAKAVDFIPVESLETTMRSP 197
                                                                                                                                                                                                                                                                                                                                                     Query Match 82.4
Best Local Similarity 77.5
Matches 158; Conservative
                                                                                                               2240
2529
2788
1035
1047
1050
1063
                                                                                                                                                                                                             1103
                                                                                              645
2041
2077
                                                                                                                                                                                                                                                                                                                                    1688
                                                                                                                                                                         1081
1085
1087
1092
1094
1097
                                                                                                                                                                  1076
                                                                                                                                                                                                                                                                                                                                           ¥.
                                                                                                                                                                                                                                                                                                                 1189
1198
1203
1680
3010 2
                                                                                                                                                            1068
                                                                                                                                                                              1082
1086
1090
                                                                                                                                                                                                 1093
                                                                                                                                                                         1077
                                                                            CARBOHYD
                                                                                             CARBOHYD
CARBOHYD
CARBOHYD
                                                                                                               CARBOHYD
CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                   STRAND
                                                                                                                                                                                                                                                                                                                                                                                                                                                           174
                   CARBOHYD
                          CARBOHYD
                                CARBOHYD
                                             CARBOHYD
                                                   CARBOHYD
                                                                     CARBOHYD
                                                                                       CARBOHYD
                                                                                                                                        HELIX
STRAND
STRAND
STRAND
TURN
STRAND
                                                                                                                                                                                                                                                  TURN
STRAND
STRAND
HELIX
TURN
                                                                                                                                                                                                                               STRAND
STRAND
STRAND
                                                                                                                                                                                                                                                                                         STRAND
                                                                                                                                                                                                                                                                                                    STRAND
                                                                                                                                                                                                              STRAND
                                                                                                                                                                                            STRAND
                                                                                                                                                                                                        STRAND
                                                                                                                                                                                                                           STRAND
                                                                                                                                                                                                                                                                                                                 STRAND
                                                                                                                                   STRAND
                                                                                                                                                                               HELIX
TURN
                                                                                                                                                                                                  TURN
                                                                                                                                                                                                                                                                                  TURN
                                                                                                                                                                                                                                                                                                                                                                                                                                           g
g
                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                    ద
                                                                                                                                                                                                                                                                                                                                                                                                                                 ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                         õ
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its most by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    MEDLINE-91192160; Pubbled-1849488;

MEDLINE-91192160; Pubbled-1849488;

MEDLINE-91192160; Pubbled-1849488;

Molkoshi S., Shimotohno K.;

Molkoshi S., Shimotohno K.;

Molecular structure of the Japanese hepatitis C viral genome.";

PEBS Lett. 280:325-328(1991).

Molecular structure of the Japanese hepatitis C viral genome.";

PEBS Lett. 280:325-328(1991).

Molecular structure of the Japanese hepatitis C viral genome.";

TEBS Lett. 280:325-328(1991).

Molecular structure of the Japanese hepatitis C viral genome.";

TEBS Lett. 280:325-328(1991).

MOSTION: THE SMALL PROFIEID NEWBRANE-RELATED FUNCTION.

MOSTALYTIC ACTIVITY: Hydrolysis of four peptide bonds in the viral precursor polyprotein, commonly with Asp or Glu in the P6 position, Cys or Thr in Pl and Ser or Ala in Pl'.

CATALYTIC ACTIVITY: N nucleoside triphosphate - N diphosphate + (RNA)(N).

CHALYTIC ACTIVITY: N nucleoside triphosphate - N diphosphate + (RNA)(N).

CHOPROTEIN ENVELOPE: THE ENVELOPE CONSISTS OF TWO PROTEINS: PROTEIN GRAND RANA.

PROTEIN CAND MRNA.

CHORDOR AND MRNA.
                                                                                                                        01-A0G-1992 (Rel. 23, Created)
01-A0G-1992 (Rel. 23, Last sequence update)
01-A0G-1992 (Rel. 24). Last sequence update)
01-A0G-1992 (Rel. 24). Last annotation update)
03-Rep-2003 (Rel. 41). Last annotation update)
Genome polyprotein [Contains: Capsid protein C (Core protein) (P22);
Envelope qlycoprotein El (GP32) (GP35); Envelope glycoprotein E2 (GP88) (GP70) (NS1); Protease/helicase NS3 (P70) (Hepaclvlrin)
(EC 3.4.22.-); Protease/helicase NS3 (P70) (Hepaclvlrin)
(EC 3.4.21.98); Nonstructural protein NS4A (P4); Nonstructural protein NS4B (P50); (NONSTRUCTURAL protein NS5A (P56); NONSTRUCTURAL protein NS5A (P66) (P70) (RNA-directed RNA Polymerase) (EC 2.7.7.48)].
Hepatitis C virus (isolate Japanese) (HCV)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          MEDLINE-91088550; PubMed-2175903;
Kato N., Hijikata M., Ootsuyama Y., Nakagawa M., Ohkoshi S., Sugimura T., Shimotohno K.;
"Molecular cloning of the human hepatitis C virus genome from Japanese patients with non-A, non-B hepatitis.";
Proc. Natl. Acad. Sci. U.S.A. 87:9524-9528(1990).
                                                                         PRT; 3010 A.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 HCV_capsid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  HCV_core.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     HCV_NS2.
HCV_NS3.
HCV_NS4a.
HCV_NS4b.
HCV_NS5a.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             EMBL; D90208; BAA14233.1; -.
                                                                     STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           IPR001490;
IPR002868;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            InterPro; IPR001410;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 IPR002522;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  nterPro; IPR002521;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   InterPro; IPR002519;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    InterPro; IPR002531;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        InterPro; IPR002518;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         IPR004109;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              IPR000745;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     A39253; GNWVCJ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              HSSP; P26663; LJXP.
MEROPS; S29.001; -.
MEROPS; U39.001; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              NCBI_TaxID=11116;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Hepacivirus
                                                                     POLG_HCVJA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             nterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              nterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               InterPro;
RESULT 5
POLG_HCVJA
                                                                                                             PARTER BERNER BE
```

```
ä
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            114 CGSSDLYLVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAV 173
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         53
                                                                                                                                                                                                                                                                                                                                  CAPSID PROTEIN ( POTENTIAL).
MATRIX PROTEIN ( POTENTIAL).
MAJOR ENVELOPE PROTEIN ( POTENTIAL).
NONSTRUCTURAL PROTEIN NS1 (POTENTIAL).
NON-STRUCTURAL PROTEIN NS2 (POTENTIAL).
PROTEASE/HELICASE NS3 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4A (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4B (POTENTIAL).
RNA-DIRECTED RNA POLYMERASE (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         3 KKGSVVIVGRINLSGD-----TAYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVST
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              54 AAQIFLAICINGVCWIVYHGAGIRIIASPKGPVIQMYINVDKDLVGWPAPQGSRSLIPCI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            9; Gaps
                                                                                                                                                                                                                                                                  Polyprotein; Glycoprotein: Transferase; RNA-directed RNA polymerase; Core protein; Coat protein: Envelope protein; Helicase; ATP-binding; Transmembrane; Nonstructural protein; Hydrolase; Serine protease.
INIT_MET

CELLGLAR AMINOPEPTIDASE.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CHARGE RELAY SYSTEM (BY SIMILARITY)
CHARGE RELAY SYSTEM (BY SIMILARITY)
CHARGE RELAY SYSTEM (BY SIMILARITY)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (POTENTIAL).
(POTENTIAL).
(POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (POTENTIAL).
(POTENTIAL).
(POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 82.4%; Score 850.5; DB 1; Length 3010; 76.5%; Pred. No. 3.8e-72; tive 23; Mismatches 16; Indels 9;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        3010 AA; 327017 MW; AA993794F46DB185 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (GLCNAC. .
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            N-LINKED (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        (GLCNAC.
(GLCNAC.
(GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  AIP (POTENTIAL)
DECH BOX.
iterPro; IPR002166; HCV_RGRP.

AterPro; IPR001650; Helicase C.

AterPro; IPR007095; RNA_DOl_DG_PS.

InterPro; IPR007094; RNA_DOl_DG_PS.

Pfam; PP01543; HCV_capsid; 1.

Pfam; PP01543; HCV_core; 1.

Pfam; PP01543; HCV_core; 1.

Pfam; PP01559; HCV_env; 1.

Pfam; PP01560; HCV_NS1; 1.

Pfam; PP01001; HCV_NS2; 1.

Pfam; PP01001; HCV_NS3; 1.

Pfam; PP01001; HCV_NS4; 1.

R Pfam; PP01001; HCV_NS4b; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                             POTENTIAL.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           N-LINKED (N-LINKED (N-LINKED (N-LINKED) (N-LINKED (N-LINKED) (N-LINKED)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      645
2041
2077
2240
2788
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Best Local Similarity
Matches 156; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            2240
2788
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1005
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ACT_SITE
ACT_SITE
NP_BIND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                             FRANSMEN
                                                                                                                                                                                                                                                                                                                                                     CHAIN
CHAIN
CHAIN
                                                                                                                                                                                                                                                                                                                                                                                            CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                      CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                    CHAIN
                                                                                                                                                                                                                                                                                                                                        CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                  CHAIN
                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ô
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMED cutstation the Buropean Bloinformatics institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                           MEDLINE-9229514; PubMed-1318627;
MEDLINE-9229514; PubMed-1318627;
MEDLINE-9229514; PubMed-1318627;
MEDLINE-9229514; PubMed-1318627;
MEDLINE-9229514; PubMed-1318627;
Medecular cloning of hepatitis C virus genome from a single Japanese cariter: sequence variation within the same individual and among infected individuals.";
Virus Res. 23:39-53(1992).
C PURCTION: THE SMALL PROTEINS NS2A, NS2B, NS4A AND NS4B ARE PURDPHOBIC, SIGGESTING A POSSIBLE MEMBRANE-RELATED FUNCTION.
NS3 AND NS5 MAY PLAY A ROLE IN THE VIRAL RNA REPLICATION.
NS3 AND NS5 MAY PLAY A ROLE IN THE VIRAL RNA REPLICATION.
C CATALYTIC ACTIVITY: Hydrolysis of four peptide bonds in the viral precursor polyprotein, commonly with Asp or Glu in the P6 possition, Cys or Thr in Pl and Ser or Ala in Pl'.
C CL- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
                                                                                                                                                     Ol-APR-1993 (Rel. 25, Created)
Ol-APR-1993 (Rel. 25, Last sequence update)
Ol-APR-1993 (Rel. 25, Last sequence update)
Ol-APR-1993 (Rel. 42, Last annotation update)
Genome polyprotein (Contains: Capsid protein C (Core protein) (P22);
Envelope diycoprotein El (GP32) (GP35); Envelope glycoprotein E2
(GP68) (GP70) (NSI); Protein P7; Nonstructural protein NS2 (P21)
(EC 3.4.21.98); Nonstructural protein NS48 (P4); Nonstructural protein
NS48 (P27); Nonstructural protein NS5A (P56); Nonstructural protein
NS5B (P60) (P70) (RNA-directed RNA polymerase) (EC 2.7.7.48)].
Hepatitis C virus (isolate HC-JT) (HCV).
Viruses; SERNA positive-strand viruses, no DNA stage; Flavivildae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      {RNA}(N).
-!- SUBMIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS: PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MRNA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           -!- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY S29.
                                                                                                                           PRT; 3010 A.A.
                 174 CTRGVAKAVDFIPVESLETIMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       RNA_pol_DS_PS.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     HCV_core.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   HCV_NS5a.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    HCV NS4b
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     HCV_env.
HCV_NS1.
HCV_NS2.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               HCV_NS4a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    EMBL; D11168; BAA01943.1; -. PIR; A45573; A45573.
                                                                                                                           STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     IPR002522;
IPR002521;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       IPR007095;
IPR007094;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            PDB; 1A10; 25-MAR-98
PDB; 1JXP; 14-JAN-98
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     InterPro; IPR001410;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               IPR000745;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         IPR002519;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    IPR001490;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        IPR002166;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                IPR004109;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              MEROPS; S29.001; -
                                                                                                                                                                                                                                                                                                                                                                                            NCBI_TaxID=31642;
                                                                                                                         POLG_HCVJT
000269;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      InterPro;
InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        InterPro;
InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (nterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            InterPro;
                                                                                                                                                                                                                                                                                                                                                                                 Hepaciv
                                                                                         RESULT
                                                                                                                                                8
```

œ

```
114 CGSSDLYLVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAV 173
                                                                                                                                                                                                               REMOVED FROM CAPSID PROTEIN C BY THE CELLULAR AMINOPEPTIDASE.
CELLULAR AMINOPEPTIDASE.
MATRIX PROTEIN (POTENTIAL).
MAJOR ENVELOPE PROTEIN E (POTENTIAL).
MAJOR ENVELOPE PROTEIN E (POTENTIAL).
NON-STRUCTURAL PROTEIN NSJ/E2 (POTENTIAL).
PROMESEK/HEICAREN RS] (POTENTIAL).
PROMESEK/HEICAREN RS] (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4A (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4A (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             3 KKGSVVIVGRIN-----LSGDTAYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVST 53
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Gaps
                                                                                                                                                                  Polyprotein; Glycoprotein; Transferase; RNA-directed RNA polymerase;
Core protein; Coat protein; Envelope protein; Helicase; ATP-binding;
Transmembrane; Nonstructural protein; Hydrolase; Serine protease;
                                                                                                                                                                                                                                                                                                                                             RNA-DIRECTED RNA POLYMERASE (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                     CHARGE RELAY SYSTEM (BY SIMILARITY).
CHARGE RELAY SYSTEM (BY SIMILARITY).
CHARGE RELAY SYSTEM (BY SIMILARITY).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                       (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         81.8%; Score 844.5; DB 1; Length 3010; 77.5%; Pred. No. 1.4e-71;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   17; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ?
·
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  94A1C77435D642BB
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                (GLCNAC.
(GLCNAC.
(GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                N-LINKED (GLCNAC.
N-LINKED (GLCNAC.
N-LINKED (GLCNAC.
N-LINKED (GLCNAC.
N-LINKED (GLCNAC.
N-LINKED (GLCNAC.
N-LINKED (GLCNAC.
N-LINKED (GLCNAC.
N-LINKED (GLCNAC.
N-LINKED (GLCNAC.
N-LINKED (GLCNAC.
N-LINKED (GLCNAC.
N-LINKED (GLCNAC.
N-LINKED (GLCNAC.
N-LINKED (GLCNAC.
N-LINKED (GLCNAC.
N-LINKED (GLCNAC.
N-LINKED (GLCNAC.
N-LINKED (GLCNAC.
N-LINKED (GLCNAC.
N-LINKED (GLCNAC.
N-LINKED (GLCNAC.
N-LINKED (GLCNAC.
N-LINKED (GLCNAC.
N-LINKED (GLCNAC.
N-LINKED (GLCNAC.
N-LINKED (GLCNAC.
N-LINKED (GLCNAC.
N-LINKED (GLCNAC.
N-LINKED (GLCNAC.
N-LINKED (GLCNAC.
N-LINKED (GLCNAC.
N-LINKED (GLCNAC.
N-LINKED (GLCNAC.
N-LINKED (GLCNAC.
N-LINKED (GLCNAC.
N-LINKED (GLCNAC.
N-LINKED (GLCNAC.
N-LINKED (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                            (GLCNAC.
(GLCNAC.
(GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Pred. No. 1.4e-71;
20; Mismatches 17
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (GICNAC.
                                                                                                                                                                                                                                                                                                                                                                                                           (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        N-LINKED (GLCNAC
                                                                                                                                                                                                                                                                                                                                                          POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CTRGVAKAVDFIPVESLEITMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        645 645 N-LIN
2041 2041 N-LIN
2240 2240 N-LIN
2529 2529 N-LIN
2788 2788 N-LIN
3010 AA: 326573 HW: C
Pfam; PF01543; HCV_capsid; 1. Pfam; PF01542; HCV_core; 1. Pfam; PF01559; HCV_core; 1. Pfam; PF01550; HCV_NS1; 1. Pfam; PF019560; HCV_NS2; 1. Pfam; PF01906; HCV_NS3; 1. Pfam; PF01006; HCV_NS4a; 1. Pfam; PF01001; HCV_NS4a; 1. Pfam; PF01001; HCV_NS5a; 1. Pfam; PF01506; HCV_NS5a; 1. Pfam; PF01506; HCV_NS5a; 1. Pfam; PF00571; helicase_C; 1. Pfam; PF00571; helicase_C; 1. Pfam; PF00998; Viral_RGRP; 1.
                                                                                                                                        ProDom; PD186062; HCV_NS1; 1.
SMART; SM00487; DEXDC; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Local Similarity 77.5
Les 158; Conservative
                                                                                                                                                                                                                                                                                                                                                          1862
2013
3010
                                                                                                                                                                                                       3D-structure.
INIT_MET
                                                                                                                                                                                                                                                                                                                                                                     ACT_SITE
ACT_SITE
ACT_SITE
NP_BIND
SITE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                               CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                   CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CARBOHYD
                                                                                                                                                                                                                                                                                                                                                         TRANSMEM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Best Loca
Matches
                                                                                                                                                                                                                                            CHAIN
CHAIN
CHAIN
                                                                                                                                                                                                                                                                              CHAIN
CHAIN
CHAIN
CHAIN
CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   셤
     à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       硆
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               à
```

```
셤
```

```
1185 CTRGVAKAVDFIPVESMETTMRSP 1208
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its was by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@lsb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Deformer Four Misser State Sta
                                                                                                                                                                            01-AUG-1992 (Rel. 23, Created)
01-AUG-1992 (Rel. 23, Last sequence update)
01-AUG-1992 (Rel. 24, Last sequence update)
01-AUG-1992 (Rel. 41, Last annotation update)
Genome polyprotein (Contains: Capsid protein C (Core protein) (P22);
Envelope dlycoprotein El (GP32) (GP35); Envelope glycoprotein E2
(GP68) (GF70) (NG1); Protean P7; Nonstructural protein NS2 (P21)
(EC 3.4.21.98); Nonstructural protein NS4A (P4); Nonstructural protein
NS4B (P27); Nonstructural protein NS5A (P56); Nonstructural protein
NS5B (P60) (P70) (RNA-dlicected RNA polymerase) (EC 2.7.7.48)].
Hepatitis C virus (isolate HC-J6) (HCV).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           (RNA)(N).
-1- SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY LIPOPROTEIN ENVELOPE CONSISTS OF TWO PROTEINS: PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MRNA
                                                                                              PRT; 3033 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      InterPro; IPR002868; HCV_NS5a.
InterPro; IPR00166; HCV_RGRP.
InterPro; IPR001650; MellCasse_C.
InterPro; IPR007095; RNL_pol_DS_PS.
InterPro; IPR007094; RNA_pol_Psvir.
Pfam; PF01543; HCV_capsid; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   HCV_capsid.
HCV_core.
HCV_env.
HCV_NSI.
HCV_NSI.
HCV_NSI.
HCV_NSI.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              EMBL; D00944; BAA00792,1; -. PIR; JQ1303; JQ1303.
                                                                                                   STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    InterPro; IPR001410;
InterPro; IPR002522;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      InterPro; IPR002519;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   InterPro; IPR000745;
InterPro; IPR001490;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     InterPro; IPR002531;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         interPro; IPR004109
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      THEI.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Hepacivirus.
NCBL_TaxID=11113;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                MEROPS; $29.001;
MEROPS; U39.001;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       PIR; JQ1303; J
HSSP; P27958;
                                                                                              POLG_HCVJ6
RESULT 7
POLG_HCVJ6
```

```
ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1034 TAYAQQTRGLLGTIVVSMTGRDKTEQAGEIQVLSTVTQSFLGTTISGVLWTVYHGAGNKT 1093
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  79 IASPKGPVIQMYTNVDKDLVGWPAPQGSRSLIPCICGSSDLYLVTRHADVIPVRRRGDSR 138
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          78
                                                                                                                                                                                                                                                                                                         CELLULAR ANINOPEPTIDASE
CAPSID PROTEIN C (POTENTIAL).
MATRIX PROTEIN (POTENTIAL).
MAJOR ENVELOPE PROTEIN E (POTENTIAL).
NONSTRUCTURAL PROTEIN NSI (POTENTIAL).
NONSTRUCTURAL PROTEIN NS2 (POTENTIAL).
PROTEASE/HELICASE NS3 (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          19 TAYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVSTAAQTFLATCINGVCWTVYHGAGTRT
                                                                                                                                                                                                                                                                                                                                                                                                                                    PROTEASE/HELICASE NS3 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4A (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4B (POTENTIAL).
RNA-DIRECTED RNA POLIMERASE (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   0; Gaps
                                                                                                                                                                                                                                   Polyprotein, Glycoprotein, Transferase, RNA-directed RNA polymerase;
Core protein, Coat protein, Envelope protein, Helicase, APP-binding;
Transmembrane; Nonstructural protein, Hydrolase; Serine protease
INIT_MET
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CHARGE RELAY SYSTEM (BY SIMILARITY).
CHARGE RELAY SYSTEM (BY SIMILARITY).
CHARGE RELAY SYSTEM (BY SIMILARITY).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (POTENTIAL).
(POTENTIAL).
(POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  (POTENTIAL). (POTENTIAL). (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            DB 1; Length 3033;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ; Pred. No. 2e-55;
26; Mismatches 31; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          N-LINKED (GLCNAC. . .)
N-LINKED (GLCNAC. . .)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     F957F5C1A273BE9E
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (GLCNAC.
(GLCNAC.
(GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   N-LINKED (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            65.3%; Score 674;
68.2%; Pred. No. 2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    POTENTIAL.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           N-LINKED
N-LINKED
N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    N-LINKED
N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                BOX.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                DECH
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     AA; 329165 MW;
                                                                                              Pfam: PF01006; HCV_NS4a; 1.
Pfam: PF01010; HCV_NS4b; 1.
Pfam: PF01010; HCV_NS4b; 1.
Pfam: PF00271; helicase_C; 1.
Pfam: PF00271; helicase_C; 1.
ProDom: P10186062; HCV_NS1; 1.
SMART; SMO0487; DEXDC; 1.
                                    HCV_NS1; 1.
HCV_NS2; 1.
HCV_NS3; 1.
  HCV_core; 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Query Match
Best Local Similarity 68.2%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1091
2038
2811
                                                                                                                                                                                                                                                                                                                                                      116
192
384
734
734
1011
1620
1867
2018
347
                                      PF01560;
PF01538;
PF01542;
                                                                             PF02907;
                    PF01539;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CHAIN
CHAIN
TRANSMEM
ACT_SITE
ACT_SITE
ACT_SITE
NP_BIND
SITE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CARBOHYD
                                      Pfam;
Pfam;
Pfam;
Pfam;
Pfam;
Pfam;
                                                                                                                                                                                                                                                                                                                                                        CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                               CHAIN
                                                                                                                                                                                                                                                                                                                                                                                               CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                        CHAIN
                                                                                                                                                                                                                                                                                                                                      CHAIN
                      Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             a
STITELLE FILLE FIL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      å
```

3033 AA

STANDARD;

POLG\_HCVJ8 P26661;

83

POLG\_HCVJ8

RESULT 8

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its way non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
01-AUG-1992 (Rel. 23, Created)
01-AUG-1992 (Rel. 23, Last sequence update)
28-FEB-2003 (Rel. 41, Last annotation update)
28-FEB-2003 (Rel. 41, Last annotation update)
Genome polyprotein [Contains: Capsid protein C (Core protein) (P22);
Envelope glycoprotein E1 (GP32) (GP35); Envelope glycoprotein E2 (GP68) (GP70) (NS1); Protein P7; Nonstructural protein NS2 (P21) (EC 3.4-22-7); Protease/helicase NS3 (P70) (Hepacivirin)
(EC 3.4-21-7); Nonstructural protein NS5A (P56); Nonstructural protein NS5B (P66); (P70) (RNA-directed RNA polymerase) (EC 2.7.7.48)].
NSSB (P65) (P70) (RNA-directed RNA polymerase) (EC 2.7.7.48)].
Vinuses; SSRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                 VITOLOGY 188:331-341(1992).

-!- FUNCTION: THE SMALL PROTEINS NSZA, NSZB, NS4A AND NS4B ARE HYDROPHOBIC, SUGGESTING A POSSIBLE MEMBRANE-RELATED FUNCTION.

NS3 AND NS5 MAY ELAY A ROLE IN THE VIRAL RNA REPLICATION.

-!- CATALYTIC ACTIVITY: Hydrolysis of four peptide bonds in the viral precursor polyprotein, commonly with Asp or Glu in the P6 position, Cys or Thr in P1 and Ser or Ala in P1'.

-!- CATALYTIC ACTIVITY: N nucleoside triphosphate - N diphosphate +
                                                                                                                                                                                                                                                                                  MEDIINE-92230232; PubMed-1314459; Okada S.-I., Yamamoto K., Lizuka H., Tanaka T., Fukuda S., Tsuda F., Mishiro S.; Tanda S., Tsuda F., Mishiro S.; Full-length sequence of a hepetitis C virus genome having poor homology to reported isolates: comparative study of four distinct
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (RNA)(N).
-1- SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLECCAPSID COVERED BY LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS: PROTEIN M AND GLYCOPROTEIN E. THE NUCLECCAPSID IS A COMPLEX OF PROTEIN C AND MRNA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY S29
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        IPR002166; HCV_RdRP.
IPR007095; RNA_pol_DS_PS.
IPR007094; RNA_pol_PSvir.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          HCV_core.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              HCV_NS4a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               HCV_NS5a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        EMBL; D10988; BAA01761.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     HCV_NS5a; 1.
Viral_RdRP;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                PF01543; HCV_caps1d;
PF01542; HCV_core; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   HCV_NS4a; 1
HCV_NS4b; 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 HCV_core; 1
HCV_env; 1.
HCV_NS1; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              InterPro; IPR007095;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          InterPro; IPR001410;
InterPro; IPR00252;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           IPR002519;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            InterPro; IPR002518;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          IPR000745;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             IPR002868;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               IPR001490;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         GNWVJ8.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         HSSP; P27958; 1HEI.
                                                                                                                                                                                                                                                                    SEQUENCE FROM N.A.
                                                                                                                                                                                                                                   NCBI_TaxID-11115;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         MEROPS; S29.001;
MEROPS; U39.001;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    PF00998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       A40250;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           InterPro;
                                                                                                                                                                                                                                                                                                                                                                      genotypes.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Pfam;
Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Pfam;
```

```
Hypothetical
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 152
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         16
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    165
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         DOMAIN
DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                HHOA_ARATH
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          LIPID
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Matches
     NAME OF THE PROPERTY OF THE PR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              d
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     79 IASPKGPVIQMYTNVDKDLVGWPAPQGSRSLTPCTCGSSDLYLVTRHADVIPVRRRGDSR 138
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    78
                                                                                            CELLULAR ANINOPEPTIDASE.
CAPSID PROTEIN C (POTENTIAL).
MAJRIX PROTEIN C (POTENTIAL).
MAJRIX PROTEIN E (POTENTIAL).
NONSTRUCTURAL PROTEIN E (POTENTIAL).
NONSTRUCTURAL PROTEIN NS2 (POTENTIAL).
NONTRUCTURAL PROTEIN NS2 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS48 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS48 (POTENTIAL).
RNA-DIRECTED RNA POLYMERASE (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   19 TAYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVSTAAQTFLATCINGVCWTVYHGAGTRT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Gaps
                                               Core protein; Coat protein; Envelope protein; Helicase; ATP-binding;
Transmembrane; Nonstructural protein; Hydrolase; Serine protease.
INIT_MET 1 1 REMOVED FROM CAPSID PROTEIN C BY THE
                 SMART; SM00487; DEXDC; 1.
Polyprotein; Glycoprotein; Transferase; RNA-directed RNA polymerase;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Treponema pallidum.
Bacteria; Spirochaetes; Spirochaetales; Spirochaetaceae: Treponema.
NCBI_TaxID=160;
                                                                                                                                                                                                                                                                           SIMILARITY).
SIMILARITY).
SIMILARITY).
                                                                                                                                                                                                                                                                                                                                                                                                                                                    (POTENTIAL)
(POTENTIAL)
(POTENTIAL)
(POTENTIAL)
(POTENTIAL)
(POTENTIAL)
(POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                        (POTENTIAL).
(POTENTIAL).
(POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                             ( POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        (POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       (POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    65.1%; Score 672; DB 1; Length 3033; 68.7%; Pred. No. 3.1e-55;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    32; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     330177 MW; 1A173E7E3381FD1A CRC64;
                                                                                                                                                                                                                                                                         CHARGE RELAY SYSTEM (BY
CHARGE RELAY SYSTEM (BY
CHARGE RELAY SYSTEM (BY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      N-LINKED (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                     N-LINKED (GLCNAC,
N-LINKED (GLCNAC,
N-LINKED (GLCNAC,
N-LINKED (GLCNAC,
N-LINKED (GLCNAC,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                      N-LINKED (GLCNAC
N-LINKED (GLCNAC
N-LINKED (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    N-LINKED (GLCNAC
N-LINKED (GLCNAC
N-LINKED (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   N-LINKED (GLCNAC
N-LINKED (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         16-0CT-2001 (Rel. 40, Created)
16-0CT-2001 (Rel. 40, Last sequence update)
16-0CT-2001 (Rel. 40, Last annotation update)
Hypothetical lipoprotein TP0136 precursor.
                                                                                                                                                                                                                                                                                                                                                           N-LINKED (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       (GLCNAC
                                                                                                                                                                                                                                                                                                                             ATP (POTENTIAL)
DECH BOX.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              485 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  24; Mismatches
                                                                                                                                                                                                                                                               POTENTIAL.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     N-LINKED (
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      N-LINKED (
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              PRT;
 PD186062; HCV_NS1; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Matches 123; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   649 64
1091 109
2038 203
2359 235
2811 281
3033 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                       423
430
448
477
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Y136_TREPA
ID Y136_TREPA
AC 083172;
                                                                                                                                                                                                                           CHAIN
CHAIN
TRANSMEM
ACT_SITE
ACT_SITE
ACT_SITE
NP_BIND
                                                                                                                                                                                                                                                                                                                                                          CARBOHYD
CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                       CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 139
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1154
                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Query Match
                                                                                                                                                                                                                                                                                                                                                                                                         CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                         CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SEQUENCE
ProDom;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Local
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            TP0136
                                                                                                              CHAIN
CHAIN
CHAIN
                                                                                                                                                                                             CHAIN
                                                                                                                                                                               CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             RESULT 9
 a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ద
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            윱
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        54 AGSKLYATNGRLWEKELNGTGSWQKVSSSSVPTDSDK ----KVMSIATDGNTFVLACVP 108
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          64 -NGVCWIVYHGAG---TRIIASPKGPVIQMYINVDKDLVG-----WPAPQGSRSLTPCT 113
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 114 C----GSSDLYLVTRHADVIP-----151
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CGGGGGGSSSSSSCIHIMLVPGGTGNNGNCGCGGGGGGSSSSSSCIHIKVENTDEQFL 224
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           109 GTGVYKHCVNGAGSSSTGTTASPSTETCSQHAT----LVGGTSKPFWLVPGGTGNNGNCG 164
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SGDTAYA-----QQTRGEEGCQETSQ----TGRDKNOVEGEVQIVSTAAQTFLATCI- 63
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          HHOA_BRATH STANDARD; PRT; 321 AA.

99SEL7; 049507;
16-0CT-2001 (Rel. 40, Created)
16-CCT-2001 (Rel. 40, Last sequence update)
15-SEP-2003 (Rel. 42, Last annotation update)
15-SEP-2003 (Rel. 42, Last annotation update)
Protease HhoA, chloroplast precursor (EC 3.4.21.-).
HHOA OR AT4G18370 OR F28312.33 o.
ATADAIOPSIS thaliana (Mouse-ear cress).
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
                                                                                    Fraser C.M., Norits S.J., Weinstock G.M., White O., Sutton G.G., Dodson R., Gwinn M., Bickey E.K., Clayton R., Ketchum K.A., Sodergren E., Hardham J.M., McLeod M.P., Salzberg S., Peterson J., Khalak H., Richardson D., Howell J.K., Chidambaram M., Utterback T., McDonald L., Artiach P., Bowman C., Cotton M.D., Fujil C., Garland S. Wenter J.C.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    71; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                              Science 281:375-388(1998).
-!- SUBCELLULAR LOCATION: Attached to the membrane by a lipid anchor
                                                                                                                                                                                                                                                                                                                                                "Complete genome sequence of Treponema pallidum, the syphilis
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            HYPOTHETICAL LIPOPROTEIN TP0136.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       (Potential)
-!- SIMILARITY: BELONGS TO THE TP013X FAMILY OF LIPOPROTEINS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Score 85.5; DB 1; Length 485;
Pred. No. 1.3;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             N-ACYL DIGLYCERIDE (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 76; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         C7A4CEEDC7DC5CED CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      protein; Lipoprotein; Membrane; Signal;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ------GSSGGPLLCPAGHAVG 167
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      225 DMGEGYVVTTKHLYTKNGSSSAGPAQCPGGGGG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                $; Pred. No. 1.3;
17; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               GLY/SER-RICH.
GLY/SER-RICH.
POLY-SER.
POLY-SER.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     GLY/SER-RICH.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    EMBL; AE001199; AAC65137.1; ALT_INIT.
                                STRAIN-Nichols;
MEDLINE-98332770; PubMed-9665876;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             48984 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               8.3%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   23.4%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             50; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      485 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Query Match
Best Local Similarity
SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Complete proteome.
```

Page 11

```
A SEAULEL FROM NA.

STRAIN-CV. Columbia.

AN MEDICAL FROM NA.

AN MEDICAL FROM NA.

AND CALL TO LOUGH FROM NA.

BOHLT. DUSSERHOEFT A. Stieken W., Entian K. D., Terryn N.,

Harris B. Ansorge W. Brandt P. O'GITUELL. Kleger M., Muchler M.,

Rechoert B., Portecell D., Perez-Alonso M., Boutry M., Bantoft I.

AND CALL MAN CALLIAGN B., Ellham L., Robben J.,

Rechoert B., Machan J., Zimmerman W., Wedler H., Ridley P.,

Langham S.-A., McCullagh B., Ellham L., Robben J.,

And Robert B., Portecell D., Perez-Alonso M., Boutry M., Barun M.,

Rechoert B., Portecell D., Perez-Alonso W., Boutry M., Barun M.,

Rechoert B., Portecell D., Perez Alonso W., Boutry M., Braun M.,

Rechoert B., Erandt A., Peters S., Van Staveren M., Dirkse W.,

Resken M., Weltjens I. Voet W., Bastiaens I., Aert R., Defoor E.,

R. Mooijaan P., Klein Lankborst R., Rose M., Hauf J., Koetter P.,

R. Mooijaan P., Klein Lankborst R., Rose M., Hauf J., Koetter P.,

R. Mooijaan P., Klein Lankborst R., Rose M., Hauf J., Koetter P.,

Bernelser S., Hempel S., Feldgausch M., Lamberth S., Van den Daele H.,

Bernelser S., Hempel S., Feldgausch M., Lennard N., McLay K., Bork R.,

R. Bockors D., Bloecker H., Scharfe M., Grimm M., Lochhert T.-H.,

Bose S., de Haam N., Mazrse A., Schaefer M., Meuller-Auer S.,

Gabel C., Fruchs M., Partmann B., Scharfe M., Grimm M., Lochhert T.-H.,

Bostova D., Bloecker H., Scharfe M., Grimm M., Lochhert T.-H.,

R. Borkova D., Bloecker H., Scharfe M., Grimm M., Lochhert T.-H.,

R. Borkova D., Bloecker R., Mearse A., Schaefer M., Mueller-Auer S.,

Gabel C., Fruchs M., Partmann B., Auburg S.,

Gabel C., Fruchs M., Partmann B., Rase D., Confeer M., Dones M., Mueller-Auer S.,

Gabel C., Fruchs M., Partmann B., Bart E., Johnson S., Tacocs P., Rechorn B.,

R. Berger C., Scholler P., Heber S., Francs P., Berkon M., Murray J., Scheet P., Cordes M., Abbort A., Scott K., Johnson D.,

R. Berchoff D., Mardie E., Johnson S., Gelsel C., Laymer B.,

R. Berchoff D., Pereston R.
                                                    SEQUENCE FROM N.A.
Lensch M.H.A., Sokolenko A., Herrmann R.G.;
"Identification and characterization of the chloroplast HhoA protease, a homolog to the bacterial periplasmic protease HhoA.";
submitted (DEC-1998) to the EMBL/GenBank/DDBJ databases.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SEQUENCE OF 72-82; 96-110; 150-159; 178-211 AND 306-320.
Schubert M., Peterson U., Funk C., Haas B., Schroeder W.
Kieselbach T.;
eurosids II; Brassicales; Brassicaceae; Arabidopsis
NCBL_TaxID-3702;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Nature 402:769-777(1999).
                                                                                                                                                                  SEQUENCE FROM N.A
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          thaliana
```

```
use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      140 ----SLLSPRPISYLK-----GSSGGPLLCPAGHAVGIFRAAVCTR---GVAKA 181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      235 SGLGREIPSPNGKSISEAIQTDADINSGNSGGPLLDSYGHTIGV-NTATFTRKGSGMSSG 293
                                                                                                                                                                                                                                                                                                                                                                                                                                                                95 KISPSVVYIEAIELPKISSGDIL----- 129
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   2 KKKGSVVIVGRINL----SGDTAYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVSTAAQT 57
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               58 FLATCINGVCW-----TVYH----TVYH------GAGTRTIASPKGPVIOM 89
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         182 -- DPDNDLAVLKIETEGRELNPVVLGTSNDLRVGQSCFAI----GNPYGYENTLTIGVV
                                                                                                                                                                                                                                                                                                                                                                                                      69; Indels 102; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          90 YINVDKDLVGWPAPQGSRSLIPCTCGSSDLYLVTRHADVIPVRRRGDSRG-----
                                                                                                                                                                                                 Chloroplast; Thylakoid; Transit peptide. CHLOROPLAST (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostoml;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NCBI_TaxID=9606;
                                                                                                                                                                                                                                                                         CHARGE RELAY SYSTEM (POTENTIAL).
CHARGE RELAY SYSTEM (POTENTIAL).
CHARGE RELAY SYSTEM (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                     8.2%; Score 85; DB 1; Length 321; 22.0%; Pred. No. 0.93;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Beckner M.E., Krutzsch H.C., Stracke M.L., Williams S.T.,
                                                                                                                                                                                                                                                                                                                           -> G (IN REF. 1).
68DB81E0BD27A7A7 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        15-JUL-1998 (Rel. 36, Created)
15-JUL-1998 (Rel. 36, Last sequence update)
28-FEB-2003 (Rel. 41, Last annotation update)
Angio-associated migratory cell protein.
                                                                                                                                                                                                                                                                                                                                                                                                   Conservative 27; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            452 AA
                                                                                                                                                                                                                                               PROTEASE HHOA.
POLY-GLU.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SEQUENCE FROM N.A., AND PARTIAL SEQUENCE
                                                                                                                                                                                                                                  THYLAKOID.
                                                                                                                                  InterPro; IPR001940; Protease2C.
InterPro; IPR001254; Ser_protease_Try.
Pfam; PF00089; trypsin; 1.
PRINTS; PR00834; PROTEASES2C.
                                                                       EMBL; AF114386; AAF24060.1;
EMBL; AL021710; CAA16717.1; ALT_SEQ.
EMBL; AL161548; CAB78839.1; ALT_SEQ.
MEROPS; S01.279;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 MEDLINE-95262124; PubMed-7743515;
                                                                                                                                                                                                                                                                                                                                        34691 MW;
                                                                                                                                                                                                    Hydrolase; Serine protease;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    182 VDF-IPVESLETTM 194
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    |:| ||:::: |:
294 VNFAIPIDTVVRTV 307
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            STANDARD;
                                                                                                                                                                                                                                                                                                                                        321 AA;
                                                                                                                                                                                                                                                                                                                                                                                      Local Similarity
nes 56; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   TISSUE-Brain;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            AAMP_HUMAN
Q13685;
                                                                                                                                                                                                                                               CHAIN
DOMAIN
ACT_SITE
                                                                                                                                                                                                                                                                                                            ACT_SITE
CONFLICT
                                                                                                                                                                                                                                                                                            _SITE
                                                                                                                                                                                                                                                                                                                                          SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                          Query Match
                                                                                                                                                                                                                 TRANSIT
                                                                                                                                                                                                                                                                                                                                                                                                        Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                RESULT 11
q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ģ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ö
```

13;

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bloinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its

Gallardo J.A., Liotta L.A.;
"Identification of a new immunoglobulin superfamily protein expressed in blood vessels with a heparin-binding consensus sequence.";
Cancer Res. 55:2140-2149(1995).
--: FUNCTION: MAY HAVE A FUNCTION IN MIGRATING CELLS.
--- TISSUE SPECIFICITY: EXPRESSED IN BLOOD VESSELS. STRONGLY EXPRESSED

```
Biol. Chem. 273:7094-7098(1998).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              152
324
171
201
280
                                                                                                                                                                                                                                MEMBRANE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Hydrolase;
TRANSIT
TRANSIT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ACT_SITE CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CONFLICT
CONFLICT
CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CONFLICT
CONFLICT
CONFLICT
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ACT_SITE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ACT_SITE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Matches
 ò
                                        This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics institute. There are no restrictions on its use by non-profil institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                   ŝ
                                                                                                                                                                                                                                                                                                                                                                                                                                        68 WIVYHGAGIRIIASPKGPVIQMYINVDKDLVGWPAPQGSRSL----TPCTCGSSDLYLV 122
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   123 TRHADVIPVRRR---GDSRGS----LLSPRPISYLKGSSG--GPLLCPA------ 162
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                241 -----VLPDGKRAVVGYEDGTIRIMDLKQGSPIHVLKGTEGHQGPLTCVAANQDGSLILT 295
                                                                                                                                                                                                                                                                                                                                                                                                                                                             197 WMEWH-----PRAPVLLAGT-ADGNTWMWKVPNGDCKTFQGPNCPATCGR----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              022609; 09LK85;
16-0CT-2001 [Rel. 40, Created)
16-0CT-2001 [Rel. 40, Last sequence update)
28-FEB-2003 (Rel. 41, Last annotation update)
Protease Do-1ike 1, chloroplast precursor (EC 3.4.21..)
DEGPI OR DEGP OR A73627925 OR K16N12.18.
Arabidopsis thallana (Mouse-ear cress).
Eukaryota: Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae:
eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 IN ENDOTHELIAL CELLS, CYTOTROPHOBLASTS, AND POORLY DIFFERENTIATED
                                                                                                                                                                                                                                                                                                                                                                                                                    Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SEQUENCE FROM N.A., AND CHARACTERIZATION.
MEDLINE-98175982: PubMed-9507020;
Itzhaki H., Naveh L., Lindahi M., Cook M., Adam Z.;
Itdentification and characterization of DegP, a serine protease associated with the luminal side of the thylakoid membrane.";
                                                                                                                                                                                                                                                                                                                                                                                                                   64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             163 ------GHAVGIFR----AAVCTRGVAKAVDFIPVESL 190
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                296 GSVDCQAKLVSATTGKVVGVFRPETVASQPSLGEGEESESNSVESL 341
                                                                                                                                                                                                                                                                                                                                                                                           7.9%; Score 82; DB 1; Length 452; 25.3%; Pred. No. 2.7;
                                                                                                                                                                                                                                                                                                                                                                                                                  13; Mismatches 47; Indels
                                                                                                                                                                                                                                                             HEPARIN-BINDING (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                        DA1413D25EB236C0 CRC64;
         COLON ADENCARCINOMA CELLS FOUND IN LYMPHATICS.
                                                                                                                                                                        GG; GG:0008201; F:heparin binding activity; TAS.
InterPro; IPR001880; Wu40.
Ffan, PF00400; WD40; 8.
SMART; SM00320; WD40; 8.
PROSITE; PS00673; WD_REPEATS_1; 1.
PROSITE; PS50082; WD_REPEATS_2; 6.
PROSITE; PS50294; WD_REPEATS_RGION; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      437 AA
                                                                                                                                                                                                                                                                         POLY - GLU
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      PRT;
                                                                                                                                                                                                                                                                                  49015 MW;
                                                                                                                              EMBL; M95627; AAA68889.1; -. PIR; I39383; I39383.
                                                                                                                                                                                                                                                                                                                                                                                                                   42; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      STANDARD;
                                                                                                                                                                                                                                                               18
77
77
138
180
220
220
261
306
363
404
                                                                                                                                                   Genew; HGNC:18; AAMP.
MIM; 603488; -.
                                                                                                                                                                                                                                                                                                                                                                                              Query Match
Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                        452 AA;
                                                                                                                                                                                                                                                    Repeat; WD repeat.
DOMAIN 14
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       DEG1_ARATH
                                                                                                                                                                                                                                                                                                                                                                        SEQUENCE
                                                                                                                                                                                                                                                                                                        REPEAT
REPEAT
                                                                                                                                                                                                                                                                          DOMAIN
                                                                                                                                                                                                                                                                                     REPEAT
                                                                                                                                                                                                                                                                                              REPEAT
                                                                                                                                                                                                                                                                                                                              REPEAT
                                                                                                                                                                                                                                                                                                                                          REPEAT
                                                                                                                                                                                                                                                                                                                                                    REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           DEG1_ARATH
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  RESULT 12
q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ŝ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ö
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ----PA 102
                                                                                                                                                                                                                                                                                         STRAIN-CV. Columbia:
Kieselbach T., Bystedt M., Schroeder W.P.;
Kleselbach T., Bystedt M., Schroeder W.P.;
Submitted (Jul-2000) to the SMISS-PROT data bank.
-!- FUNCTION: SERINE PROTEASE THAT IS REQUIRED AT HIGH TEMPERATURE.
--- HAY BE INVOLVED IN THE DEGRADATION OF DAMAGED PROTEINS. IN VIVO,
CAN DEGRADE BETA-CASEIN.
--- ENZYME REGULATION: INHIBITED BY PHENYLMETHYLSULFONYL, FLUORIDE AND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Transit peptide; Chloroplast; Thylakoid. CHLOROPLAST (POTENTIAL). THYLAKOID.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CHARGE RELAY SYSTEM (POTENTIAL).
CHARGE RELAY SYSTEM (POTENTIAL).
HSPPSSQLSNST -> SSTLFLHSPPSSHL (IN REF.
SEQUENCE FROM N.A.
STRAIN-cv. Columbia;
MEDLINE-20363099; PubMed-10907853;
KADEVO T., Katoh T., Sato S., Nakamura A., Asamizu E., Tabata S.;
Kaneko T., Katoh T., Sato S., Sato S., Seguenca A., Samizu E., Tabata S.;
Sfructural analysis of Arabidopsis thaliana chromosome 3. II.
Sequence features of the 4,251,695 bp regions covered by 90 Pl. TAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    -! - SUBCELLULAR LOCATION: BOUND TO LUMINAL SIDE OF THE THYLAKOID
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   55;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CHARGE RELAY SYSTEM (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        7.8%; Score 80.5; DB 1; Length 437; 25.6%; Pred. No. 3.5;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                55; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ---INVDKDLVGW-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 <u>--</u>
--
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1497BIABBEFFF2A4 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  2).

v -> 1 (IN REF. 2).

p -> S (IN REF. 2).

G -> R (IN REF. 2).

LL -> HF (IN REF. 2).

L -> V (IN REF. 2).

L -> V (IN REF. 2).

1 -> V (IN REF. 2).

0 -> E (IN REF. 2).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             PROTEASE DO-LIKE 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        -i- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S2C.-i- SIMILARITY: Contains 1 PDZ/DHR domain.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                SERINE PROTEASE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   18; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               70 VYHGAGIRIIASPKGPVIQMY------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           InterPro; 1901.07)
InterPro; 1PR001408; PDZ.
InterPro; 1PR001940; Protease2C.
InterPro; 1PR00154; Ser_protease_Try.
Pfam; PF00555; PDZ; 1.
Pfam; PF00089; trypsin; 1.
PRINTS; PR00884; PROTEASES2C.
SMART; SM002284; PDZ; 1.
PROSITE; PS50106; PDZ; 1.
Hydrolase; Serine protease; Transit pept
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                EMBL; AP000371; BAB02539.1; -.
EMBL; AP001302; BAB02539.1; JOINED.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 EMBL; AF028842; AAC39436.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            46213 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  -!- INDUCTION: By heat shock.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ::
<del>-</del>:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Local Similarity 25.6 es 44; Conservative
                                                                                                                                                                               and BAC clones.";
DNA Res. 7:217-221(2000).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        103
321
421
171
201
280
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              O-PHENANTHROLINE.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ..
<u>::</u>
-
                                                                                                                                                                                                                                                                      SEQUENCE OF 104-118.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                416
437 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        MEROPS; S01.279; -.
```

181 FVVARILNT 189

150 VPQGSGSGFVWDKQCHIVTNYHVIRGASDLRVTLADQTTFDAKVVGFDQDKDVAVLRIDA 209

```
CHARACTERIZATION.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                GK; Q02224;
MIM; 117143;
                                         CENE_HUMAN
                                                                                                                                                                                                                                                                       mitosis
          RESULT 14
CENE_HUMAN
                                                       This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bloinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      11;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    75 QIRVFGOND-----WMAPPASGSSAPNAMVICPCSTGTLSAVATGACNNLIERAADVALK 129
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               130 ER----RPLVLVPREAPFSSIHLENMIKLSNIGAVILPA--APGFYHQ---PQSVEDLVD 180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         86 VIQMYINVDKDLVGWPAPQGSRSLTP-----CTCGSSDL-----YLVTRHADVIPV 131
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                132 RRRGDSRGSLLSPR--PIS-----YLKGSSGGPLLCPAGHAVGIFRAAVCTRGVAKAVD 183
                          210 PK--NKLRPIPVGVSADLLVGQKVFAIGNPFGLDHTLTTGVISGLRREIS--SAATGRPI 265
103 POGSRSLTPCTCGSSDLYLV------TRHADVIPVRRGDSRGSLLSPRPI 147
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 85
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            29 QEEREVHFLISKAAQLVMAT------ETDVALPAKPQAMQAFLTEYCGAAAG 74
                                                                                                                                                                                                                                                                                                                          SEQUENCE FROM N.A.

STRAIN-ATCC 15692 / PAO1;

STRAIN-ATCC 15692 / PAO1;

STRAIN-ATCC 15692 / PAO1;

STORY C.K., Pham X.-O.T., Erwin A.L., Mizoguchi S.D., Warrener P., Stover C.K., Pham X.-O.T., Erwin A.L., Mizoguchi S.D., Warrener P., Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M., Garber R.L., Goltry L., Tolentino E., Westbrock-Wadman S., Yuan Y., Garber R.L., Coulter S.N., Folger K.R., Kas A., Larbig K., Lim R.M., Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen IT., Reizer J., Saler M.H., Hancock R.E.W., Lory S., Olson M.V.;

"Complete genome sequence of Pseudomonas aeruginosa PAO1, an
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                43 QVEGEVQ-IVSTAAQTFLATCINGVCWTVYHGACTRTIASPKGP----------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      67; Gaps
                                                       148 SYL------KGSSGGPLLCPAGHAVGIFRAAVCTRGVAKAVDF-IPVESL 190
                                                                      Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   PIR; H83144; H83144.

InterPro; IPR003382; Flavoprotein.

Pfam: PP02441; Flavoprotein; 1.

Propthetical protein; Lyase; Decarboxylase; Complete proteome.

SEQUENCE 209 AA; 22367 MW; 01PD081CC495D3F6 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                      opportunistic pathogen.";
Nature 406:959-964(2000).
-1- SIMILARITY: BELONGS TO THE POLYPRENYL P-HYDROXYBENZOATE /
PHENYLACRYLIC ACID DECARBOXYLASES FAMILY.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       7.6%; Score 78.5; DB 1; Length 209; 26.5%; Pred. No. 2.3;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Indels
                                                                                                                                                                                  16-OCT-2001 (Rel. 40, Created)
16-OCT-2001 (Rel. 40, Last sequence update)
16-OCT-2001 (Rel. 40, Last annotation update)
Probable aromatic acid decarboxylase (EC 4.1.1.-).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      26;
                                                                                                                                                        209 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Conservative 16; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       EMBL; AE004818; AAG07406.1; -.
                                                                                                                                                                                                                                                                                      Pseudomonadaceae; Pseudomonas
                                                                                                                                                        STANDARD;
                                                                                                                                                                                                                                                          Pseudomonas aeruginosa.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           184 FIPVESLET 192
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Local Similarity
hes 50; Conserv
                                                                                                                                                                                                                                                                                                    NCBI_TaxID-287;
                                                                                                                                                          PAAD_PSEAE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Best Loc
Matches
                                                                                                                                                        đ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                         a
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SEQUENCE FROM N.A.
MEDLINE-9304922; PubMed-1406971;
MEDLINE-1304922; PubMed-1406971;
FOR T.J. Li G., Schaar B.T., Szilak I., Cleveland D.W.;
*CRNP-E is a putative kinetochore motor that accumulates just before
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         -1- SUBUNIT: INTERACTS WITH CENP-F AND BUBRI KINASE.
-1- SUBCULTAR LOCATION: ASSOCIATES WITH KINETOCHORES DURING
CONGRESSION, RELOCATES TO THE SPINDLE MIDZONE AT ANAPHASE, AND
QUANTITATIVELY DISCARABED AT THE END OF THE CELL DIVISION.
-1- SIMILARITY: BELONGS TO THE KINESIN-LIKE PROTEIN FAMILY.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Thrower D.A., Jordan M.A., Schaar B.T., Yen T.J., Wilson L.; "Mitotic HeLa cells contain a CENP-E-associated minus end-directed
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CHARACTERIZATION.

CHARACTERIZATION.

CHADILIRE-848.3747. PubMed-9763420;

Chan G.K.T., Schaar B.T., Yen T.J.;

Characterization of the kinetochore binding domain of CENP-E reve

Theractions with the kinetochore proteins CENP-F and hBUBRI.";

J. Cell Biol. 143:49-63(1998).

-I- FUNCTION: MINUS-END DIRECTED MICROTUBULE MOTOR. PROBABLE

RINETOCHORE MOTOR. ACCOMULATES JUST BEFORE MITOSIS AT THE G2 F

OF THE CELL CYCLE. PROBABLY IMPORTANT FOR CHROMOSOME MOVEMENT

AND/OR SPINDLE ELONGATION.
                                                                                                                                                                                                                                                                                                                             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NCBL_TaxID=9606;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        60; GC:0005699; C:Kinetochore; TAS.
GC: GC:0005634; C:Houcleus; TAS.
GC: GC:00008350; F:Kinetochore motor activity; TAS.
GC: GC:0000067; F:Kinetochore motor activity; TAS.
GC: GC:0000067; P:DNA replication and chromosome cycle; TAS.
GC: GC:0007079; P:Hitotic chromosome movement; TAS.
GC: GC:0007089; P:Hitotic metaphase plate congression; TAS.
InterPro: IPR001752; Kinesin_motor.
                                                                                   01-JUL-1993 (Rel. 26, Created)
01-JUL-1993 (Rel. 26, Last sequence update)
30-MAY-2000 (Rel. 39, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       PRINTS; PR00380; KINESINHEAVY.
SWART; SM00129; KISC; 1.
PROSITE; PS00411; KINESIN_WOTOR_DOMAIN; 1.
PROSITE; PS50067; KINESIN_MOTOR_DOMAIN2; 1.
        2663 AA
                                                                                                                                                                                                      Centromeric protein E (CENP-E protein).
PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   MEDLINE-95196755; PubMed-7889940;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          EMBL; 215005; CAA78727.1; -. PIR; S28261; S28261.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              microtubule motor.";
EMBO J. 14:918-926(1995).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                EMD., 2.2.
PIR; S28261; S28261.
HSSP; P17119; 3KAR.
Genew; HGNC:1856; CENPE.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Pfam; PF00225; kinesin;
    STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Nature 359:536-539(1992)
                                                                                                                                                                                                                                                                                                       Homo sapiens (Human)
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                             2523 QONEQLIKQKNELLSNNQHLSNEVKTWKERTLKREAHKQVTCEN------ 2566
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     2567 SPKSPKVTGTASKKK------QITPSQCKERNL-----QDPVPKESPKSCFFD 2608
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             81 SPKGPVIQMYTNVDKDLVGWPAPQGSRSLTPCTCGSSDLYLVTRHADVIPVRRRG----D 136
                                                                                                                                                                                                                                                                                                                                                                                 32 QETSQTGRDKNQVEGEVQIVSTAAQTF-----LATCINGVCWTVYHGAGTRTIA 80
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
STRAIN—CCDG 37608 / M982 / Serogroup B / Serotype 9;
MEDLINE-9334825; PubMed=8344530;
MEDLINE-934825; PubMed=8344530;
Jacobs E., Schryvers A.B.;
                                                                                                                                                                                                                                                                                                                         Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         2609 SRSKSLPSPHPVRYFDNSSLG--LCPEVQNAG------AESVDSQP 2646
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SRG-SLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVCTRGVAKAVDFIP 186
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               -1 - FUNCTION: ACTS AS A TRANSFERRIN RECEPTOR AND IS REQUIRED FOR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      TRANSFERRIN UTILIZATION.
-1- SUBCELLULAR LOCATION: Outer membrane.
-1- INDUCTION: By iron starvation.
-1- SIMILARITY: LOCAL TO OTHER TONB-DEPENDENT RECEPTOR PROTEINS.
Motor protein; Cell division; ATP-binding; Coiled coil; Mitosis;
                                                                                                                                                                                                                                                                                                                     63;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              "Cloning and characterization of Neisseria meningitidis genes encoding the transferrin-binding proteins Tbpl and Tbp2."; Gene 130:73-80(1993).
                                                                                                                                                                                                                                                         DB 1; Length 2663;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
Neisserlaceae; Neisseria.
NCBI_TaxID=491;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   TRANSFERRIN-BINDING PROTEIN 1.
                                                                                                                                                                                                                                                                                                                     52; Indels
                                                                                                COILED COIL (POTENTIAL).
GLOBULAR (POTENTIAL).
ATP (BY SIMILARITY).
MW: CEFC13880C6C8CB8 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             38 45 TONB BOX.
894 911 TONB C-TERMINAL BOX.
911 AA; 101631 MW; 99283ABAE0B773E6 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      01-FEB-1995 (Rel. 31, Created)
01-FEB-1995 (Rel. 31, Last sequence update)
16-OCT-2001 (Rel. 40, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             911 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                PIR; JN0821; JN0821.
InterPro; IPR000531; TonB_boxC.
Pfam; PF00593; TonB_dep_Rec; 1.
PROSITE; PS00430; ToNB_DEPENBINI_REC_1; 1.
PROSITE; PS01156; TONB_DEPENBINI_REC_2; 1.
Outer membrane; Receptor: Signal; TonB_box.
                                                                                                                                                                                                                                                                                                                     15: Mismatches
                                                                                                                                                                                                                                                                                           26;
                                                                     KINESIN-MOTOR
                                                                                                                                                                                                                                                         7.5%; Score 77.5;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Transferrin-binding protein 1 precursor
                                                                                                                                                                                                                                                                                           Pred. No.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             PRT;
                                                                                                                                                                                                 2663 AA; 312087 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      EMBL; 215130; CAA78833.1; -.
                                                                                                                                                                                                                                                                                           24.08;
                                                                                                                                                                                                                                                                                                                         41; Conservative
                          Cell cycle: Centromere.
DOMAIN 1 335
DOMAIN 336 2471
DOMAIN 2472 2663
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             STANDARD;
                                                                                                                                                                   93
                                                                                                                                                                                                                                                         Ouery Match
Best Local Similarity
                                                                                                                              2472
86
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      TB11_NEIMB
Q09056;
                                                                                                                              DOMAIN
NP_BIND
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SEQUENCE
                                                                                                                                                                                                                                                                                                                            Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                STATE TO BE A READ BY A RE
      SETTT: S
                                                                                                                                                                                                                                                                                                                                                                                        ô
                                                                                                                                                                                                                                                                                                                                                                                                                                             q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
```

٠,

```
13;
                                                                                                                            :: | :|:: | | ||:
50 RRDNEVTGLGKLVKTADTLSKEQVLDIRDLTRYDPGIAVVEQGRGASSGYSIRGMDKNRV 109
                                                                                                                                                                                                                                   110 SLIVDGLAQIQSTIAQAAL-----GGTRI-AGSSGAINEIEYENVKAVEIS 154
                                                                                                                                                                                                                                                                                                                        155 ----KGSNSVEQ---GSGALAGSVAFQTKTADDVIGEGROWGIQSKTAYSGKNRGLTQS- 206
                                                                                                 43
                                                                                                                                                                                       66
                                                                                               2 KKKGSVVIVGRINLSGDTAYAQQ -----TRGEEGCQETSQ -----TGRDKNQ-
                                                                                                                                                                                       44 ---VEGEVQIVSTAAQTFLATCINGVCWTVYHGAGTRTIASPKGPVIQM-YTNVDKDLVG
                                                                                                                                                                                                                                                                               100 WPAPQGSRSLTPCTCGSSDL----YLVTRHADVIPVRR------GDSRGSLLSP
                                                   Gaps
                                                   : 99
                                                                                                                                                                                                                                                                                                                                                                         145 RPISYLKGSSGG--PLLCPAGHAVGIFRA-AVCTRGVAKAVDFIPVE 188
                                                                                                                                                                                                                                                                                                                                                                                                                ---IALAGRIGGAEALLIHTGRRAGEIRAHEDAGRGVQSFNRLVPVE 250
7.4%; Score 76; DB 1; Length 911; 26.4%; Pred. No. 22;
                                                81; Indels
                                                20; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    completed: August 30, 2003, 19:13:45
ne : 10.7567 secs
    Query Match 7.4%
Best Local Similarity 26.4%
Matches 60; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Search con
Job time
                                                                                                                                        a
                                                                                                                                                                                       ò
                                                                                                                                                                                                                                     a
                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                           ö
                                                                                                                                                                                                                                                                                   ò
```

hepatitis

091183 096146 096408 099008 099009 091174 091175 091176 091177 091177 091177 091177 091177 091177 091177 091177 091177 091177 091177 091177 091177 091177 091177 091177

091RS3 0950HG6 0950HG6 099FW19 099FW19 0991RS4 0991RS4 0991RS4 0991RR4 0991RR4 0991RR7 0991RR7 0991RR7 0991RR7 0991RR7 0991RR7 0991RR7 0991RR7 0991RR7 0991RR7

hepatitis hepatitis hepatitis hepatitis hepatitis hepatitis hepatitis

hepatitis hepatitis hepatitis hepatitis

```
SEQUENCE FROM N.A.
MEDILTB-2033484; PubMed-10864644;
MEDILTB-2033484; PubMed-10864644;
Lai V.C., Zhong W., Skelton A., Ingravallo P., Vassilev V.,
Donis R.O., Hong Z., Lau J.Y.;
"Generation and characterization of a hepatitis C virus NS3 protease-
dependent bovine viral diarrhea virus.";
J. Virol. 74:6339-6347(2000).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Lai V.C.H., Hong Z.;
Submitted (MAY-2000) to the EMBL/GenBank/DDBJ databases.
EMBL; #2268278; #ARF82566.1; -.
HSSP; P26663; LIXP.
MEROPS; S31.001; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Genome polyprotein.
Mucosal disease virus.
Viruses; ssRNA positive-strand viruses,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  MEROPS; S31.001; -.

Interpro: IPR000180; CDVII_endptseP80.

Interpro: IPR001140; DEAD.

Interpro: IPR001169; HCV_NS3.

Interpro: IPR0010569; HGV_RGRP.

Interpro: IPR0010559; HGV_BAD.

Interpro: IPR0010569; MYD_DNA_binding.

Interpro: IPR001096; RNA_pol_DS_PS.

Interpro: IPR007095; RNA_pol_DS_PS.

Interpro: IPR007094; RNA_pol_PSVIr.

Pfam; PF02907; HCV_NS3; 1.

Pfam; PF02907; HCV_NS3; 1.

Pfam; PF00998; Viral_RGRP; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    01-OCT-2000 (TYEMBLIEL 15, 01-OCT-2000 (TYEMBLIEL 15, 01-MAR-2003 (TYEMBLIEL 23,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              PRELIMINARY;
  SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               NCBI_TaxID-11099;
  889.5.5 889.5.5 889.5.5 889.5.5 889.5.5 889.5 889.5 889.5 889.5 889.5 889.5 889.5 889.5 889.5 889.5 889.5 889.5 889.5 889.5 889.5 889.5 889.5 889.5 889.5 889.5 889.5 889.5 889.5 889.5 889.5 889.5 889.5 889.5 889.5 889.5 889.5 889.5 889.5 889.5 889.5 889.5 889.5 889.5 889.5 889.5 889.5 889.5 889.5 889.5 889.5 889.5 889.5 889.5 889.5 889.5 889.5 889.5 889.5 889.5 889.5 889.5 889.5 889.5 889.5 889.5 889.5 889.5 889.5 889.5 889.5 889.5 889.5 889.5 889.5 889.5 889.5 889.5 889.5 889.5 889.5 889.5 889.5 889.5 889.5 889.5 889.5 889.5 889.5 889.5 889.5 889.5 889.5 889.5 889.5 889.5 889.5 889.5 889.5 889.5 889.5 889.5 889.5 889.5 889.5 889.5 889.5 889.5 889.5 889.5 889.5 889.5 889.5 889.5 889.5 889.5 889.5 889.5 889.5 889.5 889.5 889.5 889.5 889.5 889.5 889.5 889.5 889.5 889.5 889.5 889.5 889.5 889.5 889.5 889.5 889.5 889.5 889.5 889.5 889.5 889.5 889.5 889.5 889.5 889.5 889.5 889.5 889.5 889.5 889.5 889.5 889.5 889.5 889.5 889.5 889.5 889.5 889.5 889.5 889.5 889.5 889.5 889.5 889.5 889.5 889.5 889.5 889.5 889.5 889.5 889.5 889.5 889.5 889.5 889.5 889.5 889.5 889.5 889.5 889.5 889.5 889.5 889.5 889.5 889.5 889.5 889.5 889.5 889.5 889.5 889.5 889.5 889.5 889.5 889.5 889.5 889.5 889.5 889.5 889.5 889.5 889.5 889.5 889.5 889.5 889.5 889.5 889.5 889.5 889.5 889.5 889.5 889.5 889.5 889.5 889.5 889.5 889.5 889.5 889.5 889.5 889.5 889.5 889.5 889.5 889.5 889.5 889.5 889.5 889.5 889.5 889.5 889.5 889.5 889.5 889.5 889.5 889.5 889.5 889.5 889.5 889.5 889.5 889.5 889.5 889.5 889.5 889.5 889.5 889.5 889.5 889.5 889.5 889.5 889.5 889.5 889.5 889.5 889.5 889.5 889.5 889.5 889.5 889.5 889.5 889.5 889.5 889.5 889.5 889.5 889.5 889.5 889.5 889.5 889.5 889.5 889.5 889.5 889.5 889.5 889.5 889.5 889.5 889.5 889.5 889.5 889.5 889.5 889.5 889.5 889.5 889.5 889.5 889.5 889.5 889.5 889.5 889.5 889.5 889.5 889.5 889.5 889.5 889.5 889.5 889.5 889.5 889.5 899.5 889.5 889.5 899.5 899.5 899.5 899.5 899.5 899.5 899.5 899.5 899.5 899.5 899.5 899.5 899.5 899.5 899.5 899.5 899.5 899.5 899.5 899.5 899.5 899.5 899.5 899.5 899.5 899.5 899.5 899.5 899.5 89
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Q9IFH8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       RESULT 1
  9HJI6C
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   DDARAN DD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               091fh8 mucosal dis 081756 hepatitis c 09365 hepatitis c 0036579 hepatitis c 0931rr hepatitis c 091rr3 hepatitis c 091rr3 hepatitis c 091rr1 hepatitis c 091rr1 hepatitis c 091rr1 hepatitis c 091rr1 hepatitis c 091rr2 hepatitis c 091rr2 hepatitis c 091rr3 hepatitis c 091rr3 hepatitis c 091rr3 hepatitis c 091rr3 hepatitis c 091rr5 hepatitis c 091rr2 hepatitis c 091rr2 hepatitis c
                                                                                                                                                               (without alignments)
1352.314 Million cell updates/sec
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.
                                                                                                                                        August 30, 2003, 19:00:22; Search time 37.5921 Seconds
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Description
                                                                                                                                                                                                                   US-09-965-594-14
1032
1 MKKKGSVVIVGRINLSGDTA.....VAKAVDFIPVESLETIMRSP
                GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.
                                                                                                                                                                                                                                                                                                                                                                                                      Total number of hits satisfying chosen parameters:
                                                                                                                                                                                                                                                                                                                                                                830525 seqs, 258052604 residues
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SUMMARIES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries
                                                                                                    OM protein - protein search, using sw model
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Q91FH8
Q81756
Q91FE5
Q036579
Q91RT5
Q91RR5
Q91RR5
Q91RR5
Q91RR5
Q91RR1
Q91RQ8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Q9ELSB
Q91RR6
Q91RS9
Q91RR2
                                                                                                                                                                                                                                                                                                     BLOSUM62
Gapop 10.0 , Gapext 0.5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           sp_fung1:*
sp_human:*
sp_invertebrate:*
sp_mammal:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    sp_vertebrate:*
sp_unclassified:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          sp_rvirus:*
sp_bacteriap:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   sp_archea:*
sp_bacteria:*
sp_fungi:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         sp_organelle:*
sp_phage:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     sp_plant:*
sp_rodent:*
sp_virus:*
                                                                                                                                                                                                                                                                                                                                                                                                                                              Minimum DB seq length: 0
Maximum DB seq length: 200000000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  sp_archeap:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Match Length DB
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SPTREMBL_23:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      4040
2436
3011
3011
3011
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         181
181
181
181
181
181
181
181
181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              sp_mhc:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Query
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    910.5
901.5
901.5
900.5
896.5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 892.5
891
891
890
                                                                                                                                                                                                                          Title:
Perfect score:
                                                                                                                                                                                                                                                                                                       Scoring table:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         893
893
                                                                                                                                                                                                                                                                  Sequence:
                                                                                                                                                                                                                                                                                                                                                                  Searched:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Database
                                                                                                                                          ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Result
```

no DNA stage; Flaviviridae;

Last sequence update) Last annotation update)

Created)

PRT; 4040 AA

ALIGNMENTS

a

```
NON_TER
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             sednence.
                                                                                             Query Match
                                                                                                                                                                                                                                                                                                                                                                  O9IFE5
                                                                                                                                                                                                                                                                                                                                          RESULT
Q91FE5
                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                      οy
                                                                                                                                                                                                             a
                                                                                                                                                                                                                                     ð
                                                                                                                                                                                                                                                         셤
                                                                                                                                                                                                                                                                                  ö
                                                                                                                                                                                                                                                                                                        셤
                                                                                                                                                                                                                                                                      70 CINGVCWTVYHGAGTRIIASPKGPVIQMYTNVDQDLVGWPAPQGSRSLTPCTGGSSDLYL 129
                                                                                                                                                                                                                         62 CINGVCWTVYHGAGTRTIASPKGPVIQMYTNVDKDLVGWPAPQGSRSLTPCTCGSSDLYL 121
                                                                                                                                                                                         10 GSVVIVGRIVLSGSGSITACAQQTRGLLGCKITSLTGRDKNQVEGEVQIVSTATQTFLAT 69
                                                                                                                                                                             5 GSVVIVGRINLSGD---TAYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVSTAAQTFLAI 61
PRINTS; PR00729; CDVFNDOPTASE.
SMART; SM004047; DEXDc; 1.
PROSITE; PS00037; WYB_1; 1.
PROSITE; PS0507; RNPR_POSITIVE; 1.
PROSITE; PS0507; RNR_POSITIVE; 1.
PROSITE; PS05031; RNASE_T2_2; 1.
ATP-binding; Helicase; Hydrolase; Nonstructural protein; Polyprotein; RNA-directed RNA polymerase; Transferase.
SEQUENCE 4040 AA; 453073 WW; ADE87791D055B9DC CRC64;
                                                                                                                                                     Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae:
Hepacivirus
                                                                                                                                                     3;
                                                                                                                              Score 910.5; DB 12; Length 4040;
Pred. No. 4.5e-83;
5; Mismatches 9; Indels 3;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SEQUENCE FROM N.A.
Choo Q.-L., Richman K., Han J.;
Choo Q.-L., Richman K., Han J.;
The nucleotide sequence of the Hepatitis C viral genome.";
Submitted (MAY-1990) to the EMBL/GenBank/DDBJ databases.
EMBL; M32084; AAA45677.1; -.
HSSP; P27958; 1A1V.
                                                                                                                                                                                                                                                                                                                                                                                                 081756 PRELIMINARY; PRT; 2436 AA.
081756;
01-NOV-1996 (TrEMBLrel. 01, Created)
01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
01-NAR-2003 (TrEMBLrel. 23, Last annotation update)
denome polyprotein (Fragment).
Hepatitis C virus.
                                                                                                                              Ouery Match

88.2%; Score 910.5;
Best Local Similarity 91.3%; Pred. No. 4.5e
Matches 178; Conservative 5; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                InterPro; IPR001410; DEAD.
InterPro; IPR001531; HCV_NS.
InterPro; IPR002518; HCV_NS.
InterPro; IPR004109; HCV_NS3.
InterPro; IPR004109; HCV_NS3.
InterPro; IPR001409; HCV_NS4.
InterPro; IPR001868; HCV_NS4b.
InterPro; IPR002166; HCV_NS5a.
InterPro; IPR002166; HCV_NSP.
InterPro; IPR001059; NA_pol_DS_PS.
InterPro; IPR007094; NA_pol_DS_PS.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SMART; SM00487; DEXDC; 1.
PROSITE; PS50507; RDRP_POSITIVE; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Pfan; PF01560; HCV_NS1; I. Pfan; PF01581; HCV_NS2; I. Pfan; PF0207; HCV_NS3; I. Pfan; PF01006; HCV_NS4s; I. Pfan; PF01001; HCV_NS4b; I. Pfan; PF01506; HCV_NS5s; I. Pfan; PF00271; helicase_C; I. Pfan; PF00998; Viral_RGRP; I. Probom; PD186062; HCV_NS1; I.
                                                                                                                                                                                                                                                                                                                     182 VDFIPVESLETIMRS 196
                                                                                                                                                                                                                                                                                                                                  190 VDFIPVENLETTTRS 204
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        NCBI_TaxID-11103;
                                                                                                                                                                                                                                                                                                                                                                               RESULT
                                                                                                                                                                                                                                                                                                                                                                                                      0R
0R
0R
0R
0R
0R
0R
                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                          a
                                                                                                                                                                                                                         õ
                                                                                                                                                                                                                                               q
                                                                                                                                                                                                                                                                        οy
                                                                                                                                                                                                                                                                                              Pb
                                                                                                                                                                                                                                                                                                                      à
```

```
614
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  114 CGSSDLYLVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAV 173
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            23
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          3 KRGSVVIVGRIN-----LSGDTAYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVST
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SEQUENCE FROM N.A. MEDINE-2122; Pubmed-11369872; MEDLINE-21222; Pubmed-11369872; Lanford N.E., Lee H., Chavez D., Guerra B., Brasky K.M.; Infectious cDNA clone of the hepatitis C virus genotype 1 prototype
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    4
PROSITE; PS50521; RDRP_VIRAL; 1.
ATP-binding; Coat protein; Envelope protein; Glycoprotein; Helicase; Hydrobase; Nonstructural protein; Polyprotein;
RNA-directa RNA polymerase; Transferase; Transmembrane.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae; Hepacivirus.
NCBI_TaxID=11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       J. Geo. Virol. 82:1291-1297(2001).
-!- SUBUINT: THE VIRION OF THIS VIROS IS A NUCLEOCAPSID COVERED BY LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS: PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MRNA (BY SIMILARITY).
EMBL; AF771632; AAF81759.1; --
HSSP; P27958; 1A1V.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ŝ
                                                                                                                                                                                                                                                                                                                                                                                          87.4%; Score 901.5; DB 12; Length 2436;
85.8%; Pred. No. 1.9e-82;
tive 9; Mismatches 11; Indels 9;
                                                                                                                                                                                                                                           2436 2436
2436 AA: 264734 MW; D7B9B72900BE3125 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             01-0CT-2000 (TrEMBLrel. 15, Created)
01-0CT-2000 (TrEMBLrel. 15, Last sequence update)
01-WAR-2003 (TrEMBLrel. 23, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           PRT; 3011 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    InterPro; IPR000345; Cytc_heme_bind.
InterPro; IPR001410; DEAD.
InterPro; IPR001410; DEAD.
InterPro; IPR001522; HCV_capsid.
InterPro; IPR002519; HCV_core.
InterPro; IPR002519; HCV_env.
InterPro; IPR002519; HCV_NS1.
InterPro; IPR000749; HCV_NS2.
InterPro; IPR000749; HCV_NS4.
InterPro; IPR000749; HCV_NS4D.
InterPro; IPR001490; HCV_NS4D.
InterPro; IPR001490; HCV_NS4D.
InterPro; IPR001566; HCV_RGAD.
InterPro; IPR001666; HCV_RGAD.
InterPro; IPR001669; HGIGAD.
InterPro; IPR007094; RNA_pol_DS_FS.
INTERPRO*INTERPRO*INTERPRO*INTERPRO*INTERPRO*INTERPRO*INTERPRO*INTERPRO*INTERPRO*INTERPRO*INTERPRO*INTERPRO*INTERPRO*INTERPRO*INTERPRO*INTERPRO*INTERPRO*INTERPRO*INTERPRO*INTERPRO*INTERPRO*INTERPRO*INTERPRO*INTERPRO*INTERPRO*INTERPRO*INTERPRO*INTERPRO*INTERPRO*INTERPRO*INTERPRO*INTERPRO*INTERPRO*INTERPRO*INTERPRO*INTERPRO*INTERPRO*INTERPRO*INTERPRO*INTERPRO*INTERPRO*INTERPRO*INTERPRO*INTERPRO*INTERPRO*INTERPRO*INTERPRO*INTERPRO*INTERPRO*INTERPRO*INTERPRO*INTERPRO*INT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          174 CTRGVAKAVDFIPVESLETTMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Best Local Similarity 85.8 Matches 175; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Genome polyprotein.
Hepatitis C virus.
```

```
SEQUENCE FROM N.A.
                                       SEQUENCE FROM N.A.
                                                  STRAIN-HC-J1;
                                                           Okamoto H.;
 요
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                  Ä
                                                                                                                                                                                                                      54 AAQTFLATCINGVCWTVYHGAGTRIASPKGPVIQMYTNVDKDLVGWPAPQGSRSLTPCT 113
                                                                                                                                                                                                                                                                                        114 CGSSDLYLVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAV 173
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Okamoto H., Okada S., Sugiyama Y., Kurai K., Iizuka H., Machida A., Miyakawa Y., Mayumi M.;
"Nucleotide sequences of the genomic RNA of hepatitis C virus isolated from a human carrier: comparison with reported isolates for conserved and divergent regions.";
J. Gen. Virol. 72:2697-2704(1991).
                                                                                                                                                                                                           3 KKGSVVIVGRIN------LSGDTAYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVST
                                                                                     PROSITE; PS00190; CYTCCHROME_C; 1.
PROSITE; PS05007; RDRP_POSITIVE; 1.
ATP-Dinding: Coat protein; Envelope protein; Glycoprotein; Helicase; Hydrolase; Nonstructural protein; Potyprotein;
Hydrolase; Nonstructural protein; Potyprotein;
SEQUENCE 3011 AA; 327124 MW; 2489CE74AC864E58 CRC64;
                                                                                                                                                                                         Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                            Hepatitis C virus.
Viruses: ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                     DB 12; Length 3011;
                                                                                                                                                                                        .
6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           "The 5'-terminal sequence of the hepatitis C virus genome."; Jpn. J. Exp. Med. 60:167-177(1990).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       STRAIN-HG-JJ;
MEDLINE-91013116; PubMed-2170712;
Okamoto H.: Okada S., Sugiyama Y., Yotsumoto S., Tanaka T.,
Yoshizawa H.:
                                                                                                                                                                                         11; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                    01-NOV-1996 (TrEMBLrel. 01, Created)
01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
01-NAX-2003 (TrEMBLrel. 23, Last annotation update)
Genome polyprotein.
                                                                                                                                                                    87.4%; Score 901.5; DB 1.85.8%; Pred. No. 2.5e-82; ive 9; Mismatches 11
                                                                                                                                                                                                                                                                                                                                                                                                 PRT; 3011 AA.
                                                                                                                                                                                                                                                                                                                             STRAIN-HC-J1;
MEDLINE-92044440; PubMed-1658196;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               MEDLINE-93117120; PubMed-1335573;
Okamoto H., Kanai N., Mishiro S.;
          Pfam; PF01005; HCVAS43; 1.
Pfam; PF01001; HCVAS43; 1.
Pfam; PF01001; HCVAS54; 1.
Pfam; PF00271; helicase_C; 1.
Pfam; PF00998; Viral_RGRP; 1.
Probom; PD186062; HCVAS1; 1.
                                                                                                                                                                           Best Local Similarity 85.8
Matches 175; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                  PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        NCBI_TaxID=11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         STRAIN-HC-J1;
                                                                                                                                                                      Query Match
                                                                                                                                                                                                                                                                                                                                                                                                            003463;
                                                                                                                                                                                                                                                                                                                                                                                                  003463
Pfam;
Pfam;
Pfam;
                                                                                                                                                                                                                                                                                                                                                                               RESULT 4
                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                      a
                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                               ö
```

```
:|| :::| :::| :0005 REGREILLGPADGMVSRGWRLLAPITAYAQQTRGLIGGITSLTGRDKNQVEGEVQIVST 1064
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      114 CGSSDLYLVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAV 173
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    54 AAQIFLAICINGVCWIVYHGAGIRIIASPKGPVIQMYINVDKDLVGWPAPQGSRSLIPCI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     3 KKGSVVIVGRIN-----LSGDTAYAQQTRGEEGCQETSQTGRDVVEGEVQIVST
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SMART; SM00487; DEXDC; 1.
PROSITE; PS50507; RDRP_POSITIVE; 1.
PROSITE; PS50501; RDRP_VIRAL; 1.
ATP-binding; Coat protein; Envelope protein; Glycoprotein; Helicase;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               "Full-length nucleotide sequence of a Japanese hepatitis C virus isolate (HC-JI) with high homology to USA isolates."; Nucleic Acids Res. 20:6410-6410(1992).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      DB 12; Length 3011;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Mink M., Benichou S., Madaule P., Tiollais P., Prince A.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                12: Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Hydrolase; Nonstructural protein; Polyprotein;
RNA-directed RNA polymerase; Transferase; Transmembrane.
SEQUENCE 3011 AA; 327112 MW; 97E9052C0250463B CRC64;
                                                                                                                                                                                                                                                                                                                                                Submitted (DEC-1992) to the EMBL/GenBank/DDBJ databases
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Score 900.5; DB 1 Pred. No. 3.2e-82;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Query Match 87.3%; Score 900.5; ,
Best Local Similarity 85.8%; Pred: No. 3.2e
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Interpro; Irrolling in the control interpro; Irrolling interpro; Irrolling interpro; Irrolling interpro; Irrolling i
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         HASSY; LACADO, LALL
INTERPRO; IPRO01410;
INTERPRO; IPRO01521; HCV_capsid.
INTERPRO; IPRO02512; HCV_core.
INTERPRO; IPRO02513; HCV_env.
INTERPRO; IPRO02518; HCV_NS1.
INTERPRO; IPRO01518; HCV_NS2.
INTERPRO; IPRO01409; HCV_NS3.
INTERPRO; IPRO01409; HCV_NS3.
INTERPRO; IPRO01409; HCV_NS4.
INTERPRO; IPRO0166; HCV_NS4D.
INTERPRO; IPRO0166; HCV_NS5D.
INTERPRO; IPRO0166; HCV_NS5D.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             STRAIN-HC-J1;
MEDLINE-9417472; PubMed-7510436;
```

i,

RESULT

셤

```
54 AAQTFLATCINGVCWTVYHGAGTRTIASPKGPVIQMYTNVDKDLVGWPAPQGSRSLTPCT 113
                                                                                                                                                     114 CGSSDLYLVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAV 173
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    19 TAYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVSTAAQTFLATCINGVCWTVYHGAGTRT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               4 TAYAQQTRGLLGCIITSLTGRDKNQVEGEVQIVSTAAQTFLATCINGVCWTVYHGAGTRT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Holland-Staley C.A., Kovari L.C., Golenberg E., Mayers D.L.;
"Genetic Diversity and response to IFN of the NS3 Protease Gene from Clinical Strains of the Hepatitis C Virus.";
Submitted (APR-201) to the EMBL/GenBank/DDBJ databases.
EMBL, AF369235; AAK54560.1; -
InterPro; IPR004109; HCV_NS3.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                091RR8 PRELIMINARY; PRT; 181 AA.
091RR8;
01-DEC-2001 (TrEMBLrel. 19, Created)
01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
01-DEC-2003 (TrEMBLrel. 23, Last annotation update)
NS3 protease (Fragment).
Hepatitis C virus.
Hepatitis C virus.
Hepatitis C virus.
Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Protease.

NON_TER 1 181

NON_TER 181 181

SEQUENCE 181 AA; 19130 MW; 85D91869299B7C35 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            01-DEC-2001 (TrEMBLrel. 19, Created)
01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
01-DEC-2001 (TrEMBLrel. 23, Last annotation update)
NS3 protease (Fragment).
Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   86.8%; Score 896; DB 12;
96.6%; Pred. No. 2.5e-83;
tive 1; Mismatches 5;
                                                                                                                                                                                                                                                                     Matches 172; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SEQUENCE FROM N.A. STRAIN-Pt.4;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           NCBI_TaxID-11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  NCBI_TaxID=11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           STRAIN-Pt.1Y
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Hepacivirus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          139
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Best Local
                                                                                                                                                                                                                                                                                                                                                                                                                          RESULT 6
Q91RR8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             RESULT 7
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ACCOCCOS DE REPRESENTA DE LA PROPERCIONIO DE LA PRO
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ð
                                                                            q
                                                                                                                                                     ò
                                                                                                                                                                                                     d
                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                 a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1005 RRGQEILLGFADGMVSKGWRLLAPITAYAQQTRGLLGCIITSLTGRDKNQVEGEVQIVST 1064
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 23
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              3 KKGSVVIVGRIN-----LSGDTAYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVST
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Transmission of hepatitis C by intrahepatic inoculation with transcribed RNA.";
Science 277-570-574(1997).
-!- SUBBUT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND WRNA (BY SIMILARITY).
EMBL; AF009606; AAB66324.1; -.
HSSP; P27958; 1HEI.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  9; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            MEDLINE-97373636; PubMed-9228008;
Kolykhalov A.A., Agapov E.V., Blight K.J., Mihalik K., Feinstone S.M.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   PROSITE; PS50507; RDRP_POSITIVE; 1.
PROSITE; PS50521; RDRP_VIRAL; 1.
ATP-binding; Coat protein; Envelope protein; Glycoprotein; Helicase;
                                                                                                                                                                                                                                                                                                                                                                                                                          Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     86.9%; Score 896.5; DB 12; Length 3011;
84.8%; Pred. No. 8.2e-82;
ive 10; Mismatches 12; Indels 9;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Hydrolase; Monstructural protein; Polyprotein;
RNA-directed RNA polymerase; Transferase; Transmembrane.
SEQUENCE 3011 AA; 327182 MW; E2E0EE809C63C1B9 CRC64;
                                                                                                                                                                                                                                                                                                        Last sequence update)
Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    InterPro; IPR001410; DEAD.
InterPro; IPR00152; HCV_capsid.
InterPro; IPR00252; HCV_capsid.
InterPro; IPR00251; HCV_core.
InterPro; IPR00251; HCV_core.
InterPro; IPR00251; HCV_NS1.
InterPro; IPR001409; HCV_NS2.
InterPro; IPR001409; HCV_NS4.
InterPro; IPR001409; HCV_NS4b.
InterPro; IPR001406; HCV_NS4b.
InterPro; IPR00166; HCV_RS4b.
InterPro; IPR00166; HCV_RSAB.
InterPro; IPR00166; HCV_RSAB.
InterPro; IPR00166; HCV_RGRP.
InterPro; IPR00166; HCV_RGRP.
InterPro; IPR001694; RNA_POL_DS_PS.
InterPro; IPR010494; RNA_POL_DS_PS.
InterPro; IPR010494; RNA_POL_DS_PS.
InterPro; IPR010494; RNA_POL_DS_PS.
InterPro; IPR010494; RNA_POL_DS_PS.
                            1185 CTRGVAKAVDFIPVESLETTMRSP 1208
CTRGVAKAVDFIPVESLETTMRSP 197
                                                                                                                                                                                                                                                                              Created)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Pfam: PP01001; HCV_NS4b; 1.
Pfam: PP01506; HCV_NS5a; 1.
Pfam: PF001501; helicase_C; 1.
Pfam: PF00998; Viral_RdRP; 1.
ProDom: PD186062; HCV_NS1; 1.
                                                                                                                                                                                                                                                                        01-JAN-1998 (TrEMBLrel. 05, 01-JAN-1998 (TrEMBLrel. 05, 01-MAR-2003 (TrEMBLrel. 23,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Matches 173; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     PF01542; HCV_core;
PF01539; HCV_env; 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Pfam; PF01538; HCV_NS2; 1
Pfam; PF02907; HCV_NS3; 1
Pfam; PF01006; HCV_NS4a;
                                                                                                                                                                                                                  PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SMART; SM00487; DEXDC;
                                                                                                                                                                                                                                                                                                                                                                    Genome polyprotein. Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       NCBI_TaxID-11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              PF01560; PF01538; 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                          Hepacivirus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Rice C.M.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Query Match
                                                                                                                                                                                                                  036579
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Pfam;
```

SOW WENT TO THE PROPERTY OF TH

õ

78 63

Gaps

ö

Length 181; 5; Indels ហ

```
NCBI_TaxID=11103;
                                                                                                                                                                                                                                                                                                                                                                                                                              NCBI_TaxID-11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       STRAIN-Pt. 4B;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   STRAIN-Pt.K;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         NON_TER
NON_TER
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        NON_TER
NON_TER
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Protease.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Protease
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                RESULT 10
Q91RS1
               RESULT 9
                                                      Q91RR3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SO DE RELEGIO DE LA CONTRE LA CONTRE
                                                                                                                               Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     79 IASPKGPVIQMYTNVDKDLVGWPAPQGSRSLIPCTCGSSDLYLVTRHADVIPVRRRGDSR 138
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               64 IASPKGPVIQMYINVDKDLVGWPAPQGSRSLTPCTCGSSDLYLVTRHADVIPVRRRGDSR 123
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            79 IASPKGPVIQMYTNVDKDLVGWPAPQGSRSLTPCTCGSSDLYLVTRHADVIPVRRRGDSR 138
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         78
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                4 TAYAQQTRGLLGCIXTSLTGRDKNQVEGEVQIVSTAAQTFLATCINGVCWTVYHGAGTRT 63
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   4 TAYAQQTRGLLGCIITSLIGBDKNQVEGEVQIVSTAAQIFLAICINGVCWTVYHGAGTRI 63
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            19 TAYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVSTAAQTFLATCINGVCWTVYHGAGTRT 78
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                GSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVCTRGVAKAVDFIPVESLETTWRS 196
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              139 GSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVCTRGVAKAVDFIPVESLETTMRS 196
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  124 GSLLSPRDISYLKGSSGGPLLCPAGHAVGIFRAAVCTRGVAKAVDFIPVENLETTWRS 181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         TAYAQOTRGEEGCQETSQTGRDKNQVEGEVQIVSTAAQTFLATCINGVCWTVYHGAGTRT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Holland-Staley C.A., Kovari L.C., Golenberg E., Mayers D.L.;
"Genetic Diversity and response to IFN of the NS3 Protease Gene from
"Clinical Strains of the Hopatitis C Virus.";
Submitted (APR-2001) to the EMBL/GenBank/ODBJ databases.
EMBL, AF862538, AAK5453.1:
InterPro; IPR004109; HCV_NS3.
Holland-Staley C.A., Kovari L.C., Golenberg E., Mayers D.L.;
"Genetic Diversity and response to IFN of the NS3 Protease Gene from Clinical Strains of the Hepatitis C Virus.";
Submitted (APR-2001) to the EMBL/Genbank/DDBJ databases.
EMBL; AFS69218; AAK54543.1;
InterPro; IPR004109; HCV_NS3.
Pfam; PF02907; HCV_NS3: 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       0; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Hepatitis C virus.
Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                     Score 896; DB 12; Length 181;
Pred. No. 2.5e-83;
1; Mismatches 5; Indels (
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Query Match 86.6%; Score 894; DB 12; Length 181; Best Local Similarity 96.1%; Pred. No. 4e-83; Matches 171; Conservative 2; Mismatches 5; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         181 181
181 AA; 19084 MW; 3B5E8161F2100A72 CRC64;
                                                                                                                                                                                                                                                                                                                                                    181 181
181 AA; 19130 MW; 85D91869299B7C35 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Last sequence update)
Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   181 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  091RR5;
01-DEC-2001 (TrEMBLrel. 19, Created)
01-DEC-2001 (TrEMBLrel. 19, Last sequ
01-MAR-2003 (TrEMBLrel. 23, Last anno
NS3 protease (Fragment).
                                                                                                                                                                                                                                                                                                                                                                                                                                                             86.8%;
96.6%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Best Local Similarity 96.68 Matches 172; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Hepacivirus.
NCBI_TaxID-11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     STRAIN-Pt.3U;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       NON_TER
NON_TER
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Protease.
                                                                                                                                                                                                                                                                                   Protease.
                                                                                                                                                                                                                                                                                                                                                        NON_TER
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         13
                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Query Match
Best Local (
                                                                                                                                                                                                                                                                                                                  NON_TER
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       091RR5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         RESULT RE
```

g

ŏ

à a

g

g

ð

g

ò

ö

```
78
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             4 TAYAQOTRGLIGCIITSLIGRDKNQVBGEVQIVSTAAQIFLATCINGVCWTVYBGAGIRI 63
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            GSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVCTRGVAKAVDFIPVESLETTMRS 196
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   19 TAYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVSTAAQTFLATCINGVCWTVYHGAGTRT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Holland-Staley C.A., Kovari L.C., Golenberg E., Mayers D.L.;
"Genetic Diversity and response to IFN of the NS3 Protease Gene from Clinical Strains of the Repatitis C Virus.";
Submitted (APR-2011) to the EMBL/GenBank/DDBJ databases.
EMBL, AF369232; AAK4557.1; -.
Interpro; IPR004109; HCV_NS3; 1.
                                                                                                                                                                                                                                                                                                        Holland-Staley C.A., Kovari L.C., Golenberg E., Mayers D.L.;
"Genetic Diversity and response to IRN of the NS3 Protease Gene from
"Clinical Strains of the Hepatitis C Virus.";
Submitted (APR-2001) to the EMBL/GenBank/DDBJ databases.

EMBL, AFS62401, ARK54551;
Interpro, IPR004109; HCV_NS3.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
Hepacivirus.
                                                                                                                                             Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae; Hepacivirus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Length 181;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Length 181;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               181 181
181 AA; 19115 MW; 5D85F88AD7AC1A11 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ABB9085B3ABA4E26 CRC64;
            091RR3;
01-DEC-2001 (TrEMBLrel. 19, Created)
01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
NS3 protesse (Fragment).
Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Created)
Last sequence update)
Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   86.5%; Score 893; DB 12; 96.1%; Pred. No. 5.1e-83;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Score 893; DB 12;
Pred. No. 5.1e-83;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Pred. No. 5.1e
2; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       19114 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    86.5%;
96.1%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       01-DEC-2001 (TrEMBLrel. 19, 01-DEC-2001 (TrEMBLrel. 19, 01-MAR-2003 (TrEMBLrel. 23, MS3 protease (Fragment). Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Matches 171; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              PRELIMINARY;
PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                181 181 181 181 181 181 181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Query Match
Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                             SEQUENCE FROM N.A.
```

9

g ò

```
PRT; 3011 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Desai S.M., Devare S., Yamaguchi J.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   InterPro; IPR007095; RNA_PO1_SS_PS.
InterPro; IPR007094; RNA_PO1_SS_PS.
Pfam; PF01543; HCV_capsid; 1.
Pfam; PF01542; HCV_core; 1.
Pfam; PF01539; HCV_env; 1.
Pfam; PF01550; HCV_NSI; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                InterPro; IPR000345; Cytc_heme_bind
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           HCV_capsid.
HCV_core.
HCV_env.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Interpro; IPR000745; HCV_NS4a.
Interpro; IPR001490; HCV_NS4b.
Interpro; IPR002168; HCV_NS5a.
Interpro; IPR002166; HCV_RRPD.
Interpro; IPR002165; Helicase_C.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              InterPro; IPR001410; DEAD.
                                                                                                                                                                                                    Best Local Similarity 96.1
Matches 171; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                              PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          InterPro; IPR002521;
InterPro; IPR002519;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        InterPro; IPR002531;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    InterPro; IPR002518;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   InterPro; IPR004109;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Genome polyprotein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Hepatitis C Virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   P27958; 1HEI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           NCBI_TaxID-11103;
                                                                                                                               181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      STRAIN-colonel
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Hepacivirus
                                                                                                                               NON_TER
NON_TER
SEQUENCE
                                                                                                                 Protease.
                                                                                                                                                                                        Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                             Q9ELSB
Q9ELSB;
                                                                                                                                                                                                                                                                                                                                                                                                                                RESULT 13
Q9ELS8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                             셤
                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                쉽
                                                                                                                                                                                                                                                                                                                                                             ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     0
 ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      79 IASPKGPVIQMYINVDKDLVGWPAPQGSRSLIPCTCGSSDLYLVTRHADVIPVRRRGDSR 138
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     64 IASPKGPVIQMYTNVDKDLVGWPAPQCTRSLTPCTCGSSDLYLVTRHADVIPVRRRGDSR 123
                                                                                                  IASPKGPVIQMYTNVDKDLVGWPAPQGSRSLIPCTCGSSDLYLVTRHADVIPVRRRGDSR 138
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                19 TAYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVSTAAQTFLATCINGVCWTVYHGAGTRT 78
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               139 GSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVCTRGVAKAVDFIPVESLETTMRS 196
                                                                                                                                            GSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVCTRGVAKAVDF1PVESLETTMRS 196
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               STRAIN Pt.52;
Holland-Staley C.A., Kovari L.C., Golenberg E., Mayers D.L.;
Holland-Staley C.A., Kovari L.C., Golenberg E., Mayers D.L.;
Genetic Diversity and response to IFN of the NS3 Protease Gene from
Clinical Strains of the Hepatitis C Virus.";
Submitted (APR.2001) to the EMBL/GenBank/DDBJ databases.
EMBL; AF369245; AAK54570.1;
InterPro; IPR0044109; HCV_NS3.
Pfam; PF02907; HCV_NS3; 1.
 Gaps
                                                                                                                                                                                                                                                          091RQ8;
01-DEC-2001 (TrEMBLrel. 19, Created)
01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
01-DEC-2003 (TrEMBLrel. 23, Last annotation update)
NG3 protease (Fragment)
Hepatitis C virus
Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Hepatitis C virus.
Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
 ..
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Length 181;
 5; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   5; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             181 181
181 AA; 19144 MW; COC91F1E2EEB0B32 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           01-DEC-2001 (TIEMBLrel. 19, Created)
01-DEC-2001 (TIEMBLrel. 19, Last sequence update)
01-DEC-2003 (TIEMBLrel. 23, Last annotation update)
NS3 protease (Fragment).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     86.5%; Score 893; DB 12;
llarity 96.1%; Pred. No. 5.1e-83;
Conservative 2; Mismatches 5;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  181 AA
2; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  PRT;
Matches 171; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  PRELIMINARY;
                                                                                                                                                                                                                                              PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Similarity
                                                                                                                                                                                                                                                                                                                                                                                                     SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                NCBI_TaxID-11103;
                                                                                                                                                                                                                                                                                                                                                                          NCBI_TaxID=11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   171;
                                                                                                                                                                                                                                                                                                                                                             Hepacivirus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               NON_TER
NON_TER
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Protease.
                             13
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Query Match
Best Local
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Q91RT1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Q91RT1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Best Loca
Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    RESULT 12
091RT1
```

ga õ ò g 

```
79 IASPKGPVIQMYTNVDKDLVGWPAPQGSRSLTPCTCGSSDLYLVTRHADVIPVRRRGDSR 138
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         64 IASPKGPVIQMYINVDKDLVGWPAPQGARSLIPCTCGSSDLYLVIRHADVIPVRRGDSR 123
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   19 TAYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVSTAAQTFLATCINGVCWTVYHGAGTRT 78
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           139 GSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVCTRGVAKAVDFIPVESLETTMRS 196
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           124 GSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVCTRGVAKAVDFIPVENLETTMRS 181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Submitted (JUL-2000) to the EMBL/GenBank/DDBJ databases.
-1- SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS.
PROTEIN MAND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MRRA (BY SIMILARITY).
EMBL; AF290978; AAG02099.1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Gaps
Holland-Staley C.A., Kovari L.C., Golenberg E., Mayers D.L.;
"Genetic Diversity and response to IFN of the NS3 Protease Gene from
"Clinical Strains of the Hepatitis C Virus.";
Submitted (APR-2001) to the EMBL/GenBank/DDBJ databases.
EMBL, AFS65221, AARS6497.1;
-Interpro; IPR004109; HCV_NS3.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ö
                                                                                                                                                                                                                                                                                                                                                                                                                            Length 181;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          5; Indels
                                                                                                                                                                                                                                                                                                            181 181
181 AA: 19114 MW; ABB90B5B3ABA4E26 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            01-MAR-2001 (TrEMBLrel. 16, Created)
01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                     86.5%; Score 893; DB 12;
96.1%; Pred. No. 5.1e-83;
tive 2; Mismatches 5;
```

```
NON_TER
NON_TER
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Query Match
                                                                                                                                                                                                                                                                                                                291RS9;
                                                                                                                                                                                                                                                                                   091RS9
                                                                                                                                                                                                                                    RESULT 15
                                                                                                                                                                                                                                                              091RS9
                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  a
                                                       a
                                                                                                           ò
                                                                                                                                                          8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1125 CGSSDLYLVTRHADVIPVRRQGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGLFRAAV 1184
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          54 AAQTFLATCINGVCWTVYHGAGTRTIASPKGPVIQMYTNVDKDLVGWPAPQGSRSLTPCT 113
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        114 CGSSDLYLVTRHADVIPVRRRGDSRGSLLSPRPISYLRGSSGGPLLCPAGHAVGIFRAAV 173
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          78
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    4 TAYAQQTRGLLGCIITSLTGRDKNQVEGELQIVSTAAQTFLATCINGVCWTVFUGAGTRT 63
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         3 KKGSVVIVGRIN-----LSGDIAYAQQIRGEEGCQETSQIGRDKNQVEGEVQIVST 53
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          19 TAYAQQIRGEEGCQETSQIGRDKNQVEGEVQIVSTAAQIFLATCINGVCWIVYHGAGIRT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               STRAIN-PT.3T;
Holland-Staley C.A., Kovari L.C., Golenberg E., Mayers D.L.;
Holland-Staley C.A., Kovari L.C., Golenberg E., Mayers D.L.;
Genetic Diversity and response to Irb of the NS3 Protease Gene from Clinical Strains of the Hepatitis C Virus.";
Submitted (APR-2001) to the EMBL/GenBank/DDBJ databases.
EMBL; AR369237; AAK54562.1;
InterPro; IPR004109; HCV_NS3.
                                                                                                                                                                                                    SMART: SM00487; DEXDC: 1...
PROSITE; PS00190; CYTOCHROME_C; 1.
PROSITE; PS50507: RDRP_POSITIVE; 1.
ATP-binding; Coat protein; Envelope protein; Glycoprotein; Helicase; Hydrolase; Nonstructural protein; Polyprotein;
RNA-directed RNA polymerase; Transmembrane.
SEQUENCE 3011 AA: 327107 MW; AGBECF5A3B3EE13F CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
Hepacivirus
                                                                                                                                                                                                                                                                                                                                                                                                                                      DB 12; Length 3011;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Length 181;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        5; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     12; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             181 181
181 AA; 19101 MW; 614ADA8BOF33CCAF CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                01-DEC-2001 (TrEMBLrel. 19, Created)
01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Ouery Match

86.3%; Score 891; DB 12;
Best Local Similarity 95.5%; Pred. No. 8.2e-83;
Matches 170; Conservative 3; Mismatches 5;
                                                                                                                                                                                                                                                                                                                                                                                                                                   86.5%; Score 892.5; DB 1:
84.3%; Pred. No. 2.1e-81;
tive 11; Mismatches 12
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    181 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      3; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    174 CTRGVAKAVDFIPVESLETTMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      PRT;
Pfam; PF01538; HCV_NS2; 1.
Pfam; PF01906; HCV_NS3; 1.
Pfam; PF01006; HCV_NS4a; 1.
Pfam; PF01506; HCV_NS4b; 1.
Pfam; PF01506; HCV_NS5a; 1.
Pfam; PF00271; helicase_C; 1.
Pfam; PF00998; viral_RGRP; 1.
ProDom; PD186062; HCV_NS1; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                      Query Match
Best Local Similarity 84.3*
Matches 172; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            NS3 protease (Fragment).
Hepatitis C virus,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             NCBI_TaxID=11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   NON_TER
NON_TER
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Q91RR6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SOURCE STATE OF STATE
       DR
DR
DR
DR
DR
DR
KW
KW
SO
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      a
```

```
ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  79 IASPKGPVIQMYTHVDKDLVGWPAPQGSRSLTPCTCGSSDLYLVTRHADVIPVRRRGDSR 138
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        64 IASPRGPVIQMYINVDKDLVGWPAPQGSRSLIPCTGGSSDLYLVTRHADVIPVRRRGDSR 123
                        79 IASPKGPVIQMYTNVDKDLVGWPAPQGSRSLTPCTCGSSDLYLVTRHADVIPVRRKGDSR 138
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      78
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   63
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           139 GSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVCTRGVAKAVDFIPVESLETTMRS 196
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   139 GSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVCTRGVAKAVDFIPVESLETTMRS 196
                                                                                                                    4 TAYAQOTRGLIGCIITSLIGRDRDQVEGEVQIVSTAAQTFLATCINGVCWTVYHGAGTRT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    19 TAYAQQIRGEEGCQETSQTGRDKNQVEGEVQIVSTAAQTFLATCINGVCWTVYHGAGIRT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     "Genetic Diversity and response to IFN of the NS3 Protease Gene from Clinical Strains of the Hepatitis C Virus.";
Submitted (APR-2001) to the EMBL/Genbank/DDBJ databases.
EMBL; AF369224; AAK54549.1;
Interpro; IPR004109; BCV_NS3.
Pfam: PF02907; HCV_NS3.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                            Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
Hepacivirus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Length 181;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      181 AA; 19131 MW; 8BD7FC2769DBD635 CRC64;
                                                                                                                                                                                                                                                                                                        01-DEC-2001 (TrEMBLrel. 19, Created)
01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
01-BAR-2003 (TrEMBLrel. 23, Last annotation update)
NS3 protease (Fragment).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      86.3%; Score 891; DB 12; 96.1%; Pred. No. 8.2e-83;
                                                                                                                                                                                                                                                              181 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   2; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Search completed: August 30, 2003, 19:18:19 Job time : 38.5921 secs
                                                                                                                                                                                                                                                              PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Best Local Similarity 96.13
Matches 171; Conservative
                                                                                                                                                                                                                                                              PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                           NCBI_TaxID-11103;
                                                                                                                                                                                                                                                                                                                                                                                                           Hepatitis C virus
```

```
ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           hypothetical prote
transferrin-bindin
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    glycosyl transfera
surface-associated
chymotrypsin-like
                                                                                                                                                                                                                                                                                                                                                           genome polyprotein
                                                                                                                                                                                                                                                                                                                                                                   polyprotein
                                                                                                                                                                                                                                                                                                                                                                           polyprotein
                                                                                                                                                                                                                                                                                                                                                                                  polyprotein
polyprotein
                                                                                                                                                                                                                                                                                                                                                                                                 polyprotein
                                                                                                                                                                                                                                                                                                                                                                                                           polyprotein
                                                                                                                                                                                                                                                                                                                                                                                                                   polyprotein
polyprotein
                                                                                                                                                                                                                                                                                                                                                                                                                                  polyprotein
                                                                                                                                                                                                                                                                                                                                                                                                                                          polyprotein
                                                                                                                                                                                                                                                                                                                                                                                                                                                    - dour
                                                                                                                                                                                                                                                                                                                                                                                                                                                          polyprotein - marm
                                                                                                                                                                                                                                                                                                                                                                                                                                                                   nitrate/nitrite se
                                                                                                                                                                                                                                                                                                                                                                                                                                                                           nitrate/nitrite se
probable exported
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            prote
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    angio-associated m
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              proteinase (EC 3.4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      mosA protein - Rhi
probable periplasm
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       hypothetical prote
transferrin-bindin
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            low density lipopr
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     - Rhi
                                                        (Without alignments)
1168.492 Million cell updates/sec
                                                                                                                                                                                                                                                                                 Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            hypothetical
                                                August 30, 2003, 19:02:22; Search time 16.2134 Seconds
                                                                                                                                                                                                                                                                                                                                                                                                                                                    polyprotein
                                                                                                                                                                                                                                                                                                                                           Description
                                                                               US-09-965-594-16
1032
1 MKKKGSVVIVGRINLSGDTA.....VAKAVDFIPVESLETTMRSP
                                                                                                                                                                                                                                                                                                                                                                                                                   genome
                                                                                                                                                                                                                                                                                                                                                                                   genome
                                                                                                                                                                                                                                                                                                                                                                                                                                   genome
                                                                                                                                                                                                                                                                                                                                                                                                                                            genome
                                                                                                                                                                                                                                                                                                                                                                                                   genome
                                                                                                                                                                                                                                                                                                                                                                                                            genome
                                                                                                                                                                                                                                                                                                                                                                   genome
                                                                                                                                                                                                                                                                                                                                                                           genome
GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.
                                                                                                                                                       Total number of hits satisfying chosen parameters:
                                                                                                                                         283308 seqs, 96168682 residues
                                                                                                                                                                                                                                                                                                                  SUMMARIES
                                                                                                                                                                                                                 summaries
                                - protein search, using sw model
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            PRLJHD
B53308
B71284
                                                                                                                                                                                                                                                                                                                                                                                                        GNWVCJ
S18030
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           E75392
F81196
JE0372
                                                                                                                                                                                                                                                                                                                                                          GNWVC3
S40770
GNWVCH
GNWYTW
A45573
GNWVTC
                                                                                                                                                                                                                                                                                                                                                                                                                          JC5620
JQ1303
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1104
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   AH1030
B71360
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   139383
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    138136
                                                                                                                 BLOSUM62
Gapop 10.0 , Gapext 0.5
                                                                                                                                                                                                                                                                                                                                                                                                                                                          r08839
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    D87603
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            S61441
                                                                                                                                                                                                          1008
                                                                                                                                                                                               Post-processing: Minimum Match 0%
Maximum Match 100
Listing first 45
                                                                                                                                                                         Minimum DB seq length: 0
Maximum DB seq length: 200000000
                                                                                                                                                                                                                                                                                                                                  Query
Match Length DB
                                                                                                                                                                                                                                                                                                                                                         477
915
1615
322
2508
264
323
333
333
716
                                                                                                                                                                                                                                PIR_76:*
1: pirl:*
2: pir2:*
3: pir3:*
4: pir4:*
                                                                                                                                                                                                                                                                                                                                                                                         8883.5
878.5
867.5
8857.5
883.5
842.5
760.5
675
257
                                                                                                                 Scoring table:
                                                                                                                                                                                                                                                                                                                                          Score
                                                                                                                                                                                                                                                                                                                                                                                                                                                                 888.5
884.5
84.5
81
79
77
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   76.5
76.5
76.5
76.5
76.5
                                                                                         Perfect score:
                                 OM protein
                                                                                                  Sequence:
                                                                                                                                          Searched:
                                                                                                                                                                                                                                  Database
                                                 Run on:
                                                                                                                                                                                                                                                                                                                                                                 Title:
                                                                                                                                                                                                                                                                                                                                 Result
                                                                                                                                                                                                                                                                                                                                           ٠
ي
```

probable aromatic	hypothetical prote	hypothetical prote	hypothetical prote	centromere protein	hypothetical prote	pyruvate dehydroge	ferredoxin-nitrite	transferrin-bindin	K9R protein - vari	transferrin precur	1rreqular chiasm C	hypothetical prote	hypothetical prote	phosphodiesterase	aconitate hydratas
H83144	B83477	T04533	AB1775	S28261	T49806	DEHUPT	S16603	C81832	F72171	S33761	A49448	T35594	T16656	153865	H81775
7	~	7	~	٦	~	~	٦	7	N	~	~	~	N	7	7
209	515	846	1334	2663	354	388	594	910	318	206	764	233	429	844	898
7.3	7.3	7.3	7.3	7.3	7.3	7.3	7.3	7.3	7.2	7.2	7.2	7.2	7.2	7.2	7.2
75.5	75.5	75.5	75.5	75.5	7.5	75	75	75	74.5	74.5	74.5	74	74	74	74
30	31	32	33	34	35	36	37	38	38	0 7	41	42	43	44	4.5

## ALIGNMENTS

53

---LSGDTAYAQQTRGEEGCQETSQTGRDKNQVEGEVQLVST

3 KKGSVVIVGRIN----

```
A; Molecule type: genomic RNA
A; Residues: 1-3011 CINC.
B; Lee, D.H.; Sugitani, M.; Nasoff, M.; Prince, A.M.
B; Inchauspe, G.; Zebedee, S.; Lee, D.H.; Sugitani, M.; Nasoff, M.; Prince, A.M.
B; Inchauspe, G.; Zebedee, S.; Lee, D.H.; Sugitani, M.; Nasoff, M.; Prince, A.M.
A; Title: Genomic structure of the human prototype strain H of hepatitis C virus: comp A; Reference number: A41546; MUID:92052256; PMID:1658800
A; Contents: annotation
A; Contents: annotation
C; Superfamily: hepatitis C virus genome polyprotein
C; Superfamily: hepatitis C virus genome polyprotein
C; Reywords: ATP; capsid protein; envelope protein; glycoprotein; hydrolase; nonstruct
F; 115.7Product: envelope protein M *status predicted <RED;
F; 1907-1917Product: monstructural protein NS1 *status predicted <NS2>
F; 1307-137/Region: nucleotide-binding motif A (P-loop)
F; 1310-1317/Region: nucleotide-binding motif B
F; 1316-1319/Region: DEXH motif
F; 166-1862/Product: nonstructural protein NS4* *status predicted <N4A>
F; 183-2013/Product: nonstructural protein NS4* *status predicted <N4A>
F; 1863-2013/Product: nonstructural protein NS4* *status predicted <NAA>
F; 1813-1317/Region: DEXH motif
F; 1863-2013/Product: nonstructural protein NS4* *status predicted <NAA>
F; 1863-2013/Product: nonstructural protein NS4* *status predicted <NS5>
F; 1863-2013/Product: nonstructural protein NS4* *status predicted <NS5>
F; 1865-209, 234, 305, 325, 417, 423, 430, 448, 476, 532, 540, 556, 576, 623, 645, 1213, 12240
                                                                                                                                                                  genome polyprotein - hepatitis C virus (strain H)
N;Contains: capsid protein C; envelope protein M; hepacivirin (EC 3.4.21.98) (nonstruprotein NS4a; nonstructural protein NS4b; nonstructural protein NS5
                                                                                                                                                                                                                                                                                                                                A) Note: host Homo Sapiens (man)
C;Date: 31-Dec-1992 #sequence_revision 31-Dec-1992 #text_change 19-Jan-2001
C;Date: 31-Dec-1992 #sequence_revision 31-Dec-1992 #text_change 19-Jan-2001
C;Date: 31-Dec-1992 #sequence_revision 31-Dec-1992 #text_change 19-Jan-2001
R;Inchauspe, G.; Zebedee, S.; Lee, D.H.; Sugitani, M.; Nasoff, M.; Prince, A.M.
submitted to GenBank, July 1992
A;Reference number: A36814
A;Reference number: A36814
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Best Local Similarity 83.3%
Matches 170; Conservative
                                                                                                                                                                                                                                                                                                         Species: hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ó
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SedOvide polyprotein - hepatitis C virus
Nicottains: capsid protein C: envelope protein M; hepacivirin (EC 3.4.21.98) (nonstructured protein NS4s: nonstructural protein NS5
C:Species: hepatitis C virus
C:Accession: $40770; Pc1285
R:Comanco, H.
R:Comanco, H.
R:Comanco, H.
R:Comanco, H.: Oxada Dibrary, March 1992
A:Residues: 130170
A:Residues: 13020
A:Residues: 1-513 cok2>
A:Resid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ij
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1005 RKGREILLGPADGMYSKGWRLLAPITAYAQQTRGLLGCIITSLTGRDKNOVEGEVQIVST 1064
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1065 AAOTFLATCINGVCWTVYHGAGTRTIASPKGPVIQMYTNVDQDLVGWPAPQGARSLTPCT 1124
                                                                                                                              1125 CGSSDLYLVTRHADVIPVRRKGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAV 1184
1005 RRGREILLGPADGMYSKGWRLLAPITAYAQQTRGLLGCIITSLTGRDKNQVEGEVQIVST 1064
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    54 ATQTFLATCINGVCWTVYHGAGTRIIASPKGPVTQMYTNVDKDLVGWQAPQGSRSLTPCT 113
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               114 CGSSDLYLVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAV 173
                                                                                                                                                                                                                                                            CGSSDLYLVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAV 173
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             3 KKGSVVIVGRIN------LSGDTAYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVST
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               F;1616-1862/Product: nonstructural protein NS4a #status predicted <N4A>F;1863-2013/Product: nonstructural protein NS4b #status predicted <N4B>F;2014-3011/Product: nonstructural protein NS5 *status predicted <NS5>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      DB 1; Length 3011;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       15; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        85.6%; Score 883.5; DB 1.
llarity 84.3%; Pred. No. 8.8e-73;
Conservative 8; Mismatches 15
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1185 CTRGVAKAVDFIPVENLETTMRSP 1208
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      174 CTRGVAKAVDFIPVESLETTMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                               174 CTRGVAKAVDFIPVESLETTMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Query Match
Best Local Similarity
Matches 172; Conserv
                                                                                          54
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                )
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                              q
                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                              ð
```

```
3.4.21.98) (nonstru
                                                                                                                                               114 CGSSDLYLVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAV 173
                                                                                                                                                                                                                            54 ATQTFLATCINGVCWTVYHGAGTRTIASPKGPVTQMYTNVDKDLVGWQAPQGSRSLTPCT 113
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           genome polyprotein - hepatitis C virus (strain Taiwan)
N;Contains: capsid protein C; envelope protein M; hepacivirin (EC 3.4.21.98
protein NG4a; nonstructural protein NS4b; nonstructural protein NS5
C; Species: hepatitis C virus
A;Note: host Homo sapiens (man)
C;Date: 31.Dec-1992 *Sequence_revision 31-Dec-1992 *text_change 19-Jan-2001
C;Accession: A40244
                                                                                                                    3 KKGSVVIVGRIN------LSGDTAYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVST
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       R;Chen, P.J.; Lin, M.H.; Tai, K.F.; Liu, P.C.; Lin, C.J.; Chen, D.S.
Virology 188, 102-113, 1992
  DB 1; Length 3011;
                                                           Indels
85.1%; Score 878.5; DB 1,
83.3%; Pred. No. 2.5e-72;
tive 10; Mismatches 15;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       174 CTRGVAKAVDFIPVESLETTMRSP 197
```

us-09-965-594-16.rpr

 $\sim$ 

```
Query Match
Best Local Similarity 78.4%
Matches 160; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Matches 159; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               A; Molecule type: genomic RNA
A; Residues: 1-3010 <TAK>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Query Match
Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  C; Accession: A38465
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ø
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ô
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ö
                                                                                                                                                                                                                                                                                                                                                          a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            q
                                                                                                                                                                                                                                                                                   ð
                                                                                                                                                                                                                                                                                                                                                                                                                                 ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Genome polyprotein - hepatitis C virus (strain JT)

W. Contains: capsid protein C; envelope protein M; hepacivirin (EC 3.4.21.99) (nonstructured protein NS4s; nonstructural protein NS4b; nonstructural protein NS5
C; Species: hepatitis C virus
C; Date: 19-4My-2000 (sequence_revision 19-May-2000 *text_change 19-Jan-2001
C; Accession: A45573
R; Tanaka, T: Kato, N: Nakagawa, M: Octsuyama, Y:; Cho, M.J.; Nakazawa, T:; Hijikata, V; Tanaka, T:; Kato, N:; Nakagawa, M: Octsuyama, Y:; Cho, M.J.; Nakazawa, T:; Hijikata, V; Tanaka, T:; Kato, N:; Nakagawa, M: Octsuyama, Y:; Cho, M.J.; Nakazawa, T:; Hijikata, A; Title: Molecular cloning of hepatitis C virus genome from a single Japanese carrier: s A; Reference number: A45573; MUID: 92295714; PMID: 1318627
A; Reference number: A45573; MUID: 92295714; PMID: 1318627
A; Residuca: 1-3010 craMs
A; Molecule type: DNA
A; Mote: sequence extracted from NCEI backbone (NCEIN: 106206, NCBIP: 106207)
C; Neperfamily: hepatitis C virus genome polyprotein
C; Nepwords: ATP; glycoprotein; hydrolase; nuclectide binding; P-loop; polyprotein; F: 116-191/Product: envelope protein M *status predicted <NSI>F: 116-191/Product: monstructural protein NS1 *status predicted <NSI>F: 1107-1865/Product: monstructural protein NS1 *status predicted <NSI>F: 1107-1865/Product: nonstructural protein NS2 *status predicted <NS2)
F: 1312-1317/Region: nucleotide-binding motif A (P-loop)
F: 1312-1317/Region: nucleotide-binding motif B ** 11007-1866/Product: PREDICTED ** 11007-1866/Product: 
                                                                                             A; Molecule type: genomic RNA
A; Residues: 1.3010 <CHE>
A; Cross-references: GB:M84754
C; Superfamily: hepatitis C virus genome polyprotein
C; Superfamily: hepatitis C virus genome polyprotein
C; Superfamily: hepatitis C virus genome polyprotein; glycoprotein; hydrolase; nonstructura
C; Keywords: ATP; capsid protein C *status predicted <CPC>
F; 116-191/Product: envelope protein M *status predicted <MEE>
F; 130-1308/Product: major envelope protein B *status predicted <MEE>
F; 130-1006/Product: monstructural protein NS2 *status predicted <NS2>
F; 130-1006/Product: nonstructural protein NS2 *status predicted <NS2>
F; 1312-1317/Region: nucleotide-binding motif A (P-loop)
F; 1312-1319/Region: nucleotide-binding motif B
F; 1616-1862/Product: nonstructural protein NS4 *status predicted <N4A>
F; 166-1862/Product: nonstructural protein NS4 *status predicted <N4A>
F; 1614-3010/Product: nonstructural protein NS5 *status predicted <NAS>
F; 196, 209, 233, 234, 250, 305, 325, 417, 423, 430, 448, 532, 540, 556, 576, 623, 645, 1213, 1207
A:Title: The Taiwanese hepatitis C virus genome: sequence determination and mapping the A;Reference number: A40244; MUID:92230206; PMID:1314449
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      114 CGSSDLYLVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAV 173
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           54 ATQTFLATCINGVCWTVYHGAGTRTIASPKGPVTQMYTNVDKDLVGWQAPQGSRSLTPCT 113
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 3 KKGSVVIVGRIN------LSGDTAYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVST
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                DB 1; Length 3010;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              14; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Score 867.5; DB 1
Pred. No. 2.6e-71;
18; Mismatches 14
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    174 CTRGVAKAVDFIPVESLETTMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       84.1%;
79.9%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Matches 163; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Similarity
                                                                   A; Accession: A40244
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Local
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               A45573
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
```

```
R:Takamizawa, A.; Mori, C.; Fuke, I.; Manabe, S.; Murakami, S.; Fujita, J.; Onishi, I. J. Virol. 65, 1105-1113, 1991
A:Title: Structure and organization of the hepatitis C virus genome isolated from hur A; Reference number: A38465; MUID:91140698; PMID:1847440
A; Accession: A3846
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ACTOSS references: EMBL:MS8335; NID:g329770; PIDN:AAA72945.1; PID:g329771
C; Superfamily: hepatitis C virus genome polyprotein
C; Superfamily: hepatitis C virus genome polyprotein; dlycoprotein; hydrolase; nonstruct
C; Superfamily: hepatitis C virus genome polyprotein; dlycoprotein; hydrolase; nonstruct
C; 2-115/Product: capsid protein C *status predicted <CPC>
C; 115/Product: envelope protein N *status predicted <CPC>
C; 127-389/Product: nonstructural protein NS *status predicted <NSI>
C; 300-729/Product: nonstructural protein NS *status predicted <NSI>
C; 1007-1615/Product: hepacivirin *status predicted <NS2>
C; 1157-1137/Region: nucleotide-binding motif A (P-loop)
C; 1312-1319/Region: DEXH motif
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   F;1616-1862/Product: nonstructural protein NS4a #status predicted <N4A>
F;1863-2013/Product: nonstructural protein NS4b #status preddicted <N4B>
F;2014-3010/Product: nonstructural protein NS5 #status preddicted <NS5>
F;2014-3010/Product: nonstructural protein NS5 #status preddicted <NS5>
F;196,209,234,250,305,325,417,423,430,448,532,540,556,576,623,645,1213,1255,2041,207
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       (EC 3.4.21.98) (nonstru
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1124
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1124
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ;;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ij
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        113
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   114 CGSSDLYLVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAV 173
                                                                                                                                                                                                                                                                                                                                                                                                               113
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         53
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            genome polyprotein - hepatitis C virus
N;Contains: capsid protein C; envelope protein M; hepacivirin (EC 3.4.21.98
protein NS4a; nonstructural protein NS4b; nonstructural protein NS5
C;Species: hepatitis C virus
C;Species: hepatitis C virus
C;Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 19-Jan-2001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     3 KKGSVVIVGRIN-----LSGDTAYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVST
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             54 ATQIFLATCINGVCWTVYHGAGIRIIASPKGPVTQMYTNVDKDLVGWQAPQGSRSLIPCT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  114 CGSSDLYLVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAV
                                                                                                                                                                                                                                                                          3 KKGSVVIVGRIN-----LSGDTAYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVST
                                                                                                                                                                                                                                                                                                                                                                                                               54 ATQTFLATCINGVCWTVYHGAGTRTIASPKGPVTQMYTNVDKDLVGWQAPQGSRSLTPCT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    9; Gaps
                                                                                                                                                                                                         Gaps
F;1616-1862/Product: nonstructural protein NS4a #status predicted <N4A>F;1863-2013/Product: nonstructural protein NS4b #status predicted <N4B>F;2014-3010/Product: nonstructural protein NS5 #status predicted <NS5>
                                                                                                                                                                                                     6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Length 3010;
                                                                                                                                 Length 3010;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Indels
                                                                                                                                                                                                     Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                DB 1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    15;
                                                                                                                             83.1%; Score 857.5; DB 1;
78.4%; Pred. No. 2.2e-70;
Live 20; Mismatches 15;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            82.7%; Score 853.5; DB 1
77.9%; Pred. No. 5.1e-70;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 21; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             174 CTRGVAKAVDFIPVESLETTMRSP 197
```

```
A:Variety: isolate JKI
C;Date: 19-M97-2000 *text_change 23-Mar-2001
C;Accession: $18030; $33570; A48332; $18029
C;Accession: $18030; $33570; A48332; $18029
C;Accession: $18030
S:Honda, M.: Kaneko, S.: Massahi, U.: Kobayashi, K.: Murakami, S.
submitted to the EMEL Data Library, September 1991
A.Poescription: A whole genome of hepatitis C virus cDNA was isolated from a single pa A.Reference number: $18030
A.Rocession: $18000
A.Rocession: $18000
A.Rocession: $18000
A.Rocession: $18000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                genome polyprotein - hepatitis C virus (isolate EUH1400)
N;Contains: capsid protein C; envelope protein M; hepacivirin (EC 3.4.21.98) (nonstru protein NS4s; nonstructural protein NS4b; nonstructural protein NS5
C;Species: hepatitis C virus
C;Date: 19-May-2000 #sequence_revision 19-May-2000 #text_change 19-Jan-2001
C;Accession: JC5620
S;Chansolan, R.W.; Adams, N.J.; Taylor, L.A.; Simmonds, P.; Elliott, R.M.
Biochem: Blophys. Res Commun. 236, 44-49, 1997
A;Title: The complete coding sequence of hepatitis C virus genotype 5a, the predomina A;Reference number: JC5620; MUID:97366593; PMID:9223423
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        F;1616-1862/Product: nonstructural protein NS4m *status predicted <N4A>
F;1863-2013/Product: nonstructural protein NS4b *status predicted <N4B>
F;2014-3010/Product: nonstructural protein NS5 *status predicted <NS5>
F;2014-3010/Product: nonstructural protein NS5 *status predicted <NS5>
F;196,209,234,250,305,417,423,448,532,540,556,576,623,645/Binding site: carbohydrate
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1005 RRGREILLGPADGFREQGWRLLAPITAYSQQTRGLFGCIVTSLTGRDKNQVEGEAQVVST 1064
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       173
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             114 CGSSDLYLVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               54 ATQTFLATCINGVCWTVYHGAGTRTIASPKGPVTQMYTNVDKDLVGWQAPQGSRSLTPCT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     3 KKGSVVIVGRIN-----LSGDIAYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVSI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      81.6%; Score 842.5; DB 1; Length 3010; llarity 77.5%; Pred. No. 5.3e-69; Conservative 20; Mismatches 17; Indels 9;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CIRGVAKAVDFIPVESLETIMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         al Similarity
158; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               174
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1185
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Best Local
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                         Cypecies: hepatitis C virus
Cypecies: hydroxics constructural protein No.50 function 1902 *text_change 19-Jan-2001
CyAccession: A3953; ps6086
A.Tille: Wolecular choning of the human hepatitis C virus genome from Japanese patients
A.Tille: Molecular choning of the human hepatitis C virus genome from Japanese patients
A.Reference number: A3953; MUTD:9108850; PMID:2175903
A.Reference consider the factor of the human hepatitis viral genome show sequence varia
A.Reference signification of the non-A, non-B hepatitis viral genome show sequence varia
A.Reference number: PS0085
A.Accession: PS0086
A.Accession: PS0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    genome polyprotein - hepatitis C virus (isolate JKl)
N:Contains: capsid protein C; envelope protein M; hepacivirin (EC 3.4.21.98) (nonstructu protein NS4a; nonstructural protein NS4b; nonstructural protein NS5 C; Species: hepatitis C virus
                                                                                                                                                                                                                                                                                                                        genome polyprotein - hepatitis C virus (strain J)
N;Contains: capsid protein C; envelope protein M; major envelope protein E; nonstructura
protein NS4a; nonstructural protein NS4b; nonstructural protein NS5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                114 CGSSDLYLVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAV 173
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   54 ATQTFLATCINGVCWTVYHGAGIRTIASPRGPVTQMYTNVDKDLVGWQAPQGSRSLTPCT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         3 KKGSVVIVGRINLSGD-----TAYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVST
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   82.7%; Score 853.5; DB 1; Length 3010; 77.0%; Pred. No. 5.1e-70;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          15; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          23; Mismatches
                                                           1185 CTRGVAKAVDFIPVESMETTMRSP 1208
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CTRGVAKAVDFIPVESLETIMRSP 197
                  CTRGVAKAVDFIPVESLETTMRSP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Query Match
Best Local Similarity
Matches 157; Conserv
```

g à 셤

g à

õ

```
65.6%;
68.7%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Query Match
Best Local Similarity 69.3%.
Matches 124; Conservative
                                                                                                                                                                                                            Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       F;1320-1323/Region: DEXH motif
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         A; Molecule type: genomic RNA A; Residues: 1-3033 <OKA>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            A; Molecule type: mRNA
A; Residues: 2678-2729 <KAT>
                                                                                                                                                             Similarity
                                                                                                                                                                                                  Matches 123;
                                                                                                                    Query Match
Best Local
                                                                                                                                                                                                                                                                                                                                                                          a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       å
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   A. Coramoto, H.; Okada, S.; Sugiyama, Y.; Kurai, K.; Iizuka, H.; Machida, A.; Miyakawa, Y. J. Gen. Virol. 72, 2697-2704, 1991
A.71tler Nucleotide sequence of the genomic RNA of hepatitis C virus isolated from a hum A.Reference number: 301303; MUID:92044440; PMID:1658196
A.Accession: J01303
A.Moleoule type: genomic RNA
A.Residues: 1-3033 <OKA>
A.Coss.references: GB:D00944; NID:9221650; PIDN:BAA00792.1; PID:9221651
A.Experimental Source: isolate Hc-J6 from a Japanese individual
C.Superfamily: hepatitis C virus genome polyprotein
C.Superfamily: Pepatitis C virus genome polyprotein; serine proteinase; transm
F:2-115/Product: capsid protein C *status predicted <CPC>
F:116-119/Product: major envelope protein M *status predicted <MEE>
F:390-733/Product: monstructural protein NS1 *status predicted <NS2>
F:1316-1321/Region: nonstructural protein NS2 *status predicted <NS2>
F:1316-1321/Region: DEXH motif
A; Residues: 1-3014 <CHA>
A; Residues: 1-3014 <CHA>
A; Residues: 1-3014 <CHA>
A; Residues: 1-3014 <CHA>
A; Cross_references: Genotype 5a, which predominates in South Africa
A; Cross_references: Genotype 5a, which predominates in South Africa
A; Cross_references: Genotype 5a, which predominates in not complete in this paper
C; Superfamily: hepatitis C virus genome polyprotein
C; Superfamily: hepatitis C ** Status predicted c CPC>
F; 130-130/Product: major envelope protein B ** status predicted c NS2>
F; 131-130/Product: nonstructural protein NS1 ** status predicted c NS2>
F; 1317-1320/Region: nucleotide-binding motif A (P-loop)
F; 1317-1320/Region: DEXH motif
F; 1317-1320/Region: DEXH motif
F; 1317-1320/Region: nucleotide-binding motif B
F; 1317-1320/Region: nucleotide-binding B
F; 
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               genome polyprotein - hepatitis C virus (isolate HC-J6)
N;Contains: capsid protein C; envelope protein M; hepacivirin (EC 3.4.21.98) (nonstructu protein NS4s; nonstructural protein NS4b; nonstructural protein NS5 (species : hepatitis C virus
C;Date: 19-May-2000 %sequence_revision 19-May-2000 #text_change 17-Nov-2000
C;Accession: JQ1303
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1006 RRGREIFLGPADDIKTSGWRLLAPITAYAQQTRGVLGAIVLSLTGRDKNEAEGEVQFLST 1065
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Ä
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 54 ATQIFLATCINGVCWTVYHGAGIRIIASPKGPVTQMYINVDKDLVGWQAPQGSRSLIPCI 113
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             114 CGSSDLYLVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAV 173
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     3 KKGSVVIVGRIN-----LSGDTAYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVST
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     F;1620-1866/Product: nonstructural protein NS4a #status predicted <N4A>F;1867-2017/Product: nonstructural protein NS4b #status predicted <N4B>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Ouery Match

73.7%; Score 760.5; DB 1; Length 3014;
Best Local Similarity 69.6%; Pred. No. 1.9e-61;
Matches 142; Conservative 24; Mismatches 29; Indela o.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1186 CTRGVAKALEFVPVENLETTMRSP 1209
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  174 CTRGVAKAVDFIPVESLETTMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             õ
```

```
ACCOSS-TREFERENCES: GB.D90518; NID:9221523; PIDN:BAA01418.1; PID:9221524

ACCOSS-TREFERENCES: GB.D90518; NID:9221523; PIDN:BAA01418.1; PID:9221524

CSUPErfamily: hepatitis C virus genome polyprotein
CSUPERIALLY: papsid protein: envelope protein: glycoprotein; hydrolase; nonstruc
E:1-115/Product: capsid protein C *status predicted <CPC>
E:116-191/Product: envelope protein M *status predicted <MEE>
E:192-1894/Product: major envelope protein M *status predicted <MEE>
E:330-733/Product: nonstructural protein NS1 *status predicted <NS1>
E:101-1619/Product: nonstructural protein NS2 *status predicted <NS2>
E:101-1619/Product: hepatolylrin *status predicted <NS2>
E:1214-1211/Region: nucleotide-binding motif A (P-loop)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                genome polyprotein - hepatitis C virus (strain HC-J8)
N;Contains: capsid protein C; envelope protein M; hepacivirin (EC 3.4.21.98) (nonstrested in Ns4a: nonstructural protein NS4a: nonstructural protein NS5
C;Species: hepatitis C virus
C;Species: hepatitis C virus
C;Date: 31-Dec-1992 #sequence_revision 31-Dec-1992 #text_change 19-Jan-2001
C;Accession: A40250; pp0397; pp0559
R;Okamoto, H.; Kurai, K.; Okada, S.I.; Yamamoto, K.; Lizuka, H.; Tanaka, T.; Fukuda, Virology 188, 331-341, 1992
A;Title: Full-length sequence of a hepatitis C virus genome having poor homology to A;Reference number: A40250; MUID:92230232; PMID:1314459
F;2018-3033/Product: nonstructural protein NS5 #status predicted <NO5>
F;196,209,234,305,325,417,423,430,448,477,534,542,558,578,627,649,1091,1217,1259,203
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                A: Molecule type: genomic RNA
A: Residues: 2678-2754 CCRN>
A: Cross-references: DBJ:DJ0134
A: Experimental source: isolate E-bl2
R: Kato, N.; Ootsuyama, Y.; Ohkoshi, S.; Nakazawa, T.; Mori, S.; Hijikata, M.; Shimot
Biochem. Biophys. Res. Commun. 181, 279-285, 1991
A; Title: Distribution of plural HCV types in Japan.
A; Reference number: P00554; MUID:92068204; PMID:1720309
A; Accession: P00559
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         F;1620-1866/Product: nonstructural protein NS4a *status predicted <N4A>
F;1867-2017/Product: nonstructural protein NS4b *status predicted <N4B>
F;2018-3033/Product: nonstructural protein NS5 *status predicted <NS5>
F;2018-303,299,305,417,423,430,448,477,534,542,558,578,627,649,1091,1217,1259,203
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          A;Cross-references: GB:D10988; GB:D01221; NID:g221608; PIDN:BAA01761.1; PID:g221609 R;Chan, S.W.; McOmish, F.; Holmes, E.C.; Dow, B.; Peutherer, J.F.; Follett, E.; Yap, J. Gen. Virol. 73, 1131-1141, 1992
A;Title: Analysis of a new hepatitis C virus type and its phylogenetic relationship A;Reference number: PQ0393; WUID:92268871; PMID:1316939
                                                                                                                                                                                                                               ö
                                                                                                                                                                                                                                                                                                                                                                   ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     79 IASPKGPVTQMYTNVDKDLYGWQAPQGSRSLTPCTCGSSDLYLVTRHADVIPVRRRGDSR 138
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       139 GSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVCTRGVAKAVDFIPVESLETTMRSP 197
                                                                                                                                                                                                                                                                                                        19 TAYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVSTATOTFLATCINGVCWTVYHGAGTRT
                                                                                                                                                                                                                                    Gaps
                                                                                                                                                                                                                          ;
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         0:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         65.4%; Score 675; DB 1; Length 3033;
69.3%; Pred. No. 1.5e-53;
.1ve 24; Mismatches 31; Indels
                                                                                                                                                                                                                          30;
                                                                                                                                    .6%; Score 677; DB 1; .7%; Pred. No. 9.8e-54;
                                                                                                                                                                                                                      26; Mismatches
```

Oy 53 TATOTFLATCINGVCWTVYHGAGTRTIASI  :   :  :  :  :  :  :  :  :  :  :  :  :	RESULT 14 B81104 nitrate/nitrite sensor protein (EC 2.7.3) C.Species: Neisseria meningitidis C.Species: Neisseria meningitidis C.Species: Neisseria meningitidis C.Species: Neisseria meningitidis C.Species: Neisseria Nitratesion 31-Mc C.Scession: B81104 R.Tettelin, H.; Saunders, N.J.; Heidelberg, Hickey, ER, Haft, D. H.; Salzberg, S.L.; Mi ri, H.; Oin, H.; Vamathevan, J.; Gill, J.; Science 287, 1809-1815, 2000 A.Tuthors: Grandl, G.; Sun, L.; Salth, H.O., A.Title: Complete genome sequence of Neissen A.Reference number: A81000; MUID:20175755; IA:Molecule type: DNA	A. Kesidues: 1-590 <tet> A. Cross-references: GB:AE002473; GB:AE002098 A. Experimental source: serogroup B, strain b C. Genetics: A. Gene: NMB1249 C. Superimally: nitrate/nitrite sensor prote: (Seywords: autophosphorylation: phosphohist F;395/Active site: His (Phosphohistidine in</tet>	Query Match Best Local S Matches 47	28 213	OY  BE OMTINION SHOOL OF SHOOL OF SHOOL OF SHOOL SHOOL OF SHOOL SHOOL OF SH	lantematical. Language
OY 19 TAYAOOTRCEEGCOETSOTGRDKNQVEGEVQIVSTATQTFLATCINGVCWTVYHGAGTRT 78	RESULT 12 T08841  POLYPOTED - douroucouli hepatitis GB virus A C;Species: 20.56p-1999 #sequence_revision 20.5ep-1999 #text_change 17.Nov-2000 C;Accession: T08841 B;Erker, J.C.; Desai, S.M.; Leary, T.P.; Chalmers, M.L.; Montes, C.C.; Mushahwar, I.K. A;Title: Genomic analysis of two GB virus A variants isolated from captive monkeys. A;Reference number: 216486; MUID:98120818; PMID:9460920 A;Accession: T08841 A;Status: translated from GB/EMBL/DDBJ A;Status: translated from GB/EMBL/DDBJ A;Status: translated from GB/EMBL/DDBJ A;Status: translated from GB/EMBL/DDBJ A;Coss-references: EMBL:AF023425; NID:92828599; PIDN:AAC40502.1; PID:928286600 C;Superfamily: hepatitis C virus genome polyprotein C;Keywords: polyprotein	Query Match  24.9%; Score 257; DB 2; Length 3005;  Best Local Similarity 34.1%; Pred. No. 5.2e-15;  Matches 56; Conservative 29; Mismatches 69; Indels 10; Gaps 3;  Qy 33 ETSQTGRDKNQVECEVQIVSTATGTFLATCINGVCMTVYHGAGTRTIASPKGPVTQMTN 92  1	93 VDKDLVGAQARQGSRSLTPCTCGSSDLYLVT   :                   : : :   1055 PSDDVAVYPLPSGASCLEPCKCGTQSVWCIR	QY L53 SGGFLLCPSGHANGTRANCTRGYAKAVDFIPYES 189 	RESULT 13 T108839 POLYPOTOLEIN - marmoset hepatitis GB virus A C; Species: marmoset hepatitis GB virus A C; Sacession: T08839 A; Gen. VIrol. 79, 41-45, 1998 A; Title: Genomic analysis of two GB virus A variants isolated from captive monkeys. A; Reference number: 216486; MUID:98120818; PMID:9460920 A; Reference number: 216486; MUID:98120818; PMID:9460920 A; Accession: T08839 A; Accession: T08839 A; Molecule type: genomic RNA A; Residues: 1-2970 cERK> A; Molecule type: genomic RNA A; Cross-references: EMBL:AF023424; NID:92828597; PIDN:AAC40501.1; PID:92828598 C; Superfamily: hepatitis C virus genome polyprotein C; Reywords: polyprotein C; Reywords: polyprotein Cuery Match Best Local Similarity 27.8%; Pred. No. 1.8e-14; Matches 62; Conservative 39; Mismatches 80; Indels 42; Gaps 6; A; Conservative 39; Mismatches 80; Indels 42; Gaps 6; A; H	

```
Bentley, S.D.; Churcher, C.; Klee, S.R.; Monule, S.; Mungall, K.; Quail, M.A.; Rajandre
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ..; Fraser, C.M.; Moxon, E.R.; Rappuoll, R.;
eria meningitidis serogroup B strain MC58.
PMID:10710307
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ein narX
stidine; phosphoprotein; phosphotransferase
intermediate) *status predicted
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    NMA1418 [similarity] - Neisseria meningit
                                                                                                                                                                                                                                                                                                                             ) NMB1249 [similarity] - Neisseria meningit
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             198; NID:97226488; PIDN:AAF41629.1; PID:9722
MC58
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              oup A strain of Neisseria menigitidis 22491
PMID:10761919
                                                                                                                                                                                                                                                                                                                                                                                                                   , J.; Jeffries, A.C.; Nelson, K.E.; Eisen, White, O.; Fleischmann, R.D.; Dougherty, B. Scarlato, V.; Masignani, V.; Pizza, M.
                              : :| :|
ECGVA-----EQTRSLEKONQNLT 259
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ---GSRSLTPCTCGSSDLYLVTRHAD---- 127
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                |: | | | | ||:||: ||||
PAVGADSGRVCLDGGSDVYVSIHHADCGTAAS 318
SPKGPVTQMYTNVDKDLVGWQAPQGSRSLTPC 112
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SYLKGSSGGPLLCPAGHAVGIFRAA----- 172
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    LLS------PRPISTLKGSSGGPLLCP 161
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             OTFLATCINGVCWTVYHGAGTRTIASPKGPVT 87
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1.5; DB 2; Length 590;
0. 2.7;
ttches 76; Indels 65; Gaps
                                                                                                                                                                                                                                                                                                                                                                         Mar-2000 *text_change 19-Jan-2001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  May-2000 *text_change 02-Feb-2001
                                                                                                                                                                                                    : :| |:: | |
ETLPREAITHTEAPP 1154
                                                                                                                                                                            DFIPVESLETTMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            - IPVESLET 192
: |: || || ||
NLOVOMLET 414
```

```
A; Molecule type: DNA
A; Residues: 1-590 <PAR>
A; Residues: 1-590 <PAR>
A; Residues: 1-590 <PAR>
A; Cross=**references: GB:AL162755; GB:AL157959; NID:g7379742; PIDN:CAB84658.1; PID:g738007
A; Experimental source: serogroup A, strain 22491
C; Genetics: MA1418
A; Gene: MA1418
C; Superfamily: nttrate/nitrite sonsor protein narx
C; Reywords: autophosphorylation; phosphohistidine; phosphoprotein; phosphotransferase; s
F; 395/Active site: His (phosphohistidine intermediate) **status predicted*

A; 295/Active site: His (phosphohistidine intermediate) **status predicted*
                                                                                                                                                                                                                                                                                                                                                                                                                          88 QMYTNVDKDLVGWQAPQ-----127
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            128 -----VIPVRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAA----- 172
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 319 DLGKYHEBIFPIEYQNETLGRILLSFPNGISLDEDDRILLQTLGRQLGVSLAGAKQEEEK 378
                                                                                                                                                                                                                                                                                                                                                                                        28 EEGCQETSQTGRDKNQVEGEVQIVSTATQTFLATCINGVCWTVYHGAGTRTIASPKGPVT 87
                                                                                                                                                                                                                                                                                                                                        65; Gaps
                                                                                                                                                                                                                                                                            Length 590;
                                                                                                                                                                                                                                                                            Query Match 8.6%; Score 88.5; DB 2; Length 59
Best Local Similarity 21.8%; Pred. No. 2.7;
Matches 47; Conservative 28; Mismatches 76; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  173 ------VCTRG----VAKAVDF--IPVESLET 192
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    379 RLLAVLQERNLIAQGLHDSIAQALTFLNLQVQMLET 414
                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ολ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
```

Search completed: August 30, 2003, 19:20:29 Job time: 18.2134 secs

Run on:

```
ARI18696 Sequence 109331 Sequence 15 ARI18708 Sequence ARI1872 Sequence ARI1872 Sequence ARI1872 Sequence BD091382 HCV cultiples CARIS91 Sequence 10 ARIS91 Sequence ARIA6483 Sequence ARIA6483 Sequence ARIA6483 Sequence 9 I81885 Sequence 9 M62321 Hepatitis CARIS980 Synthetic ARIA7806 Synthetic ARIA7806 Synthetic EREA59 Synthetic EREA59 Synthetic EREA59 Synthetic EREA56 CONA Of Hep
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    D10749 Repatitis C
AX395309 Sequence
AX454818 Sequence
AX127810 Sequence
BD081911 Hepatitis
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      AF511948 Heparitis
AR127809 Sequence
BD081910 Heparitis
AR10828 Sequence
BD06992 Functiona
AF09606 Heparitis
AR110831 Sequence
BD069985 Functiona
                                                                                                                                                                                                                                    AR179057 Sequence
AF268278 Pestiviru
AR118686 Sequence
106434 Sequence 48
109328 Sequence 8
                                                                                                                                                                                                                                                                                                 106440 Sequence 54
109329 Sequence 10
                                                                                                                                                                                                                                                                                        AR118692 Sequence
                                                                                                                                           Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.
                                                                                                                                                                                                                    Description
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ALIGNMENTS
                                                                                                                                                                                   SUMMARIES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     HPCPLYPRE
AF387805
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       AF387808
AF271632
AF387806
AF387807
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       AF511948
AR127809
                                                                                                                                                                                                                                                                                                                                                                                                    108294
AR166930
AR301300
AR176483
BD080334
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             AX395309
AX454818
AR127810
                                                                                                                                                                                                                                                                                                                                           HPCPOLYP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           AF009606
                                                                                                                                                                                                                                                                                                                                                                                AR118723
BD091382
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                E08263
E08264
HPCHCJ1
                                                                                                                                                                                                                                                                                                                                                    AR118703
                                                                                                                                                                                                                                                                                                                                                                        AR118722
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             BD081910
                                                                                                                                                                                                                                                                                                                                                               AR118728
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                BD081911
                                                                                                                                                                                                                                                                                                                                                                                                                                                         E66293
        em_htg_hum:*
em_htg_inv:*
em_htg_other:*
                                                                                                                       em_htgo_other:*
                                                                                           em_sy:*
em_htgo_hum:*
                                                                                em_htg_vrt:*
                                        em_htg_mus:*
                                                  em_htg_pln:*
em_htg_rod:*
                                                                     em_htg_mam:*
                                                                                                              em_htgo_mus
                                                                                                                                                                                                      Query
Match Length DB
                                                                                                                                                                                                                                                                                                 RESULT 1
                                                                                                                                                                                                                                                                     Result
                                                                                                                                                                                                                  Š.
                                                                                                                                                                                                                                                                                                                            Command line parameters:
-WODEL-frame+-p2n.model -DEV-x1p
-WODEL-frame+-p2n.model -DEV-x1p
-WODEL-frame+-p2n.model -DEV-x1p
-Word 2_1/GSF0_spool/GS09965594/runat_29082003_151919_28310/app_query.fasta_1.2872
-GJ-cqn2_1/GSF0_spool/GS09965594/runat_29082003_151919_28310/app_query.fasta_1.2872
-DB-GenEmbl .OFMT-fastap -SUFPIX-rege -MINMARCH-0.1 -LOOPCL-0 -LOOPEXT-0
-UNITS-b1ts .START-1 -EMD-1 - MARTAX-blosum62 -TRANS-human40.cd1 -11ST-45
-DCCALIGN-200 -THR_SCORE-pct -THR_MAX-100 -THR_MIN-0 -ALIGN-15 -WODE-LOCAL
-USTR-US9965594_eCGN _1_114686_etunat_29082003_151919_28310 -NCPU-6 -ICPU-3
-NO_MMAP -LARGEQUERY -NEG_SCORES-0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_INMEQUT-120 -WARN TIMEOUT-30 -THREADS-1 -XGAPOP-10 -XGAPEXT-0.5 -FGAPOP-6
-FGAPEXT-7 -XGAPOP-10 -YGAPEXT-0.5 -DELEXT-7
                                                                             (without alignments)
3147.423 Million cell updates/sec
                                                                                                                       1032
1 MKKKGSVVIVGRINLSGDTA......VAKAVDFIPVESLETTMRSP 197
                                                                    August 30, 2003, 19:18:33 ; Search time 2560.57 Seconds
      GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.
                                                 OM protein - nucleic search, using frame_plus_p2n model
                                                                                                                                                                                                                2888711 segs, 20454813386 residues
                                                                                                                                                                                                                                     of hits satisfying chosen parameters:
                                                                                                                                                                                                                                                                                     Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries
                                                                                                                                                 BLOSUM62
Xgapop 10.0 , Ygapext (
Ygapop 10.0 , Ygapext (
Ygapop 6.0 , Fgapext Delop 6.0 , Delext
                                                                                                                                                                                                                                                        Minimum DB seq length: 0
Maximum DB seq length: 200000000
                                                                                                            US-09-965-594-14
                                                                                                                                                                                                                                                                                                                                                                                                                                                            gb_ba:*
gb_htg:*
gb_om:*
gb_ow:*
gb_pat:*
gb_ph:*
gb_ph:*
gb_pr:*
gb_ro:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       em_fun:*
em_hum:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              em_pat:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             em_ro:*
em_sts:*
em_un:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             gb_sy:*
gb_un:*
gb_vi:*
em_ba:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             em_in:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     em_mu:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 еш_оm: *
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          em_or: *
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     em_ov:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   em_pl:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                        GenEmbl:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Perfect score:
Sequence:
                                                                                                                                                      Scoring table:
                                                                                                                                                                                                                                      rotal number
                                                                                                                                                                                                                                                                                                                                                                                                                                                       Database :
                                                                                                                                                                                                                Searched:
```

~

```
LLISEDLPAAVRNIMARTDHPEPIQLAYNSYEVQVPVLFPKIRNGEYTDTYENYSFLN
ARKLGEDVPVYIYATEDEDLAVDLIGLDMPDPCNQOVVETGRALKQYTGLSSAENALL
                                                                                                                                                                                                                                 Location/Qualifiers
  AF268278
AF268278.1 GI:9049956
                                    type 1
                                                                                                                                                                                                                                                                                      385
                                    Pestivirus
                                                Pestivirus
                                                                                                                                                              10864644
                                                                                                                                                                                                                                           source
                                                                                                                                                                                                                                                                                     5'UTR
CDS
                                               ORGANISM
                                                                                                                                                  MEDLINE
PUBMED
  ACCESSION
             VERSION
KEYWORDS
SOURCE
                                                                                REFERENCE
                                                                                           AUTHORS
                                                                                                                                                                        REFERENCE
                                                                                                                                                                                   AUTHORS
                                                                                                                                                                                                          JOURNAL
                                                                                                                                       JOURNAL
                                                                                                                                                                                                                               FEATURES
                                                                                                                TITLE
                                                                                                                                                                                               TITLE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    VRL 12-JUL-2000
              PAT 20-APR-2002
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        113 GGTAGTGTTGTTATTGTTGGTAGAATTGTTTTATCTGGTAGTGGTAGTATCACGCCGTAC 472
                                                                                                                                                                                                                                                                                                                                                                                         652
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ProLysGlyProValileGlnMetTyrThrAsnValAspLysAspLeuValGlyTrpPro 101
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CCCAAGGGTCCTGTCATCCAGATGTATACCAATGTGGACCAAGACCTTGTGGGCTGGCCC 712
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            102 AlaProGlnGlySerArgSerLeuThrProCysThrCysGlySerSerAspLeuTyrLeu 121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        LeuSerProArgProlleSerTyrLeuLysGlySerSerGlyGlyProLeuLeuCysPro 161
                                                                                                                                                                                                                                                                                                                                                                                                                                                                     81
                                                                                                                                                                                                                                                                                                                                                                              41
                                                                                                                                                                                                                                                                                                                                                                                                                          61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                     GlySerValvalileValGlyArgileAsnLeuSerGlyAsp-----ThrAlaTyr
                                                                                                                                                                                                                                                                                                                                                                             AlaGlnGlnThrArgGlyGluGluGlyCysGlnGluThrSerGlnThrGlyArgAspLys
                                                                                                                                                                                                                                                                                                                                                                                                                          AsnGlnValGluGlyGluValGlnIleValSerThrAlaAlaGlnThrPheLeuAlaThr
                                                                                                  1 (bases 1 to 12734)
Hong, Z., Lai, W.C.H. and Lau, J.Y.N.
Hong, Z., Cai, W.C.H. and Lau, J.Y.N.
Hepatitis C virus protease-dependent chimeric pestivirus
Patent: US 6326137-A 1 04-DEC-2001;
Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              182 ValAspPheIleProValGluSerLeuGluThrThrMetArgSer 196
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          linear
              linear
                                                                                                                                                                                                                           Length:
Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    RNA
               DNA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  AF268278 12734 bp Pestivirus type 1, complete genome.
                                                                                                                                                                                   2803
        Sequence 1 from patent US 6326137.
AR179057
                                                                                                                                                                                                                                                                                                           (1-12734)
                                                                                                                                                                     /organism="unknown"
a 2604 c 3295 9
                                                                                                                                                                                                                                                                                                           US-09-965-594-14 (1-197) x AR179057
                                             AR179057.1 GI:20220612
                                                                                                                                                                                                                              6.27e-65
930.50
94.87%
92.82%
                                                                                           Unclassified
                                                                                                                                                                                                                                                                 Similarity:
                                                                                                                                                                                   4032
                                                                      Unknown
                                                                                  Unknown
                                                                                                                                                                                                                                                      Percent Similarity:
                                                                                                                                                                                                                    Aliqnment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         122
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      142
                                                                                                                                                                                                                                                                                                                                                                                                   173
                                                                                                                                                                                                                                                                                                                                                                                                                                                333
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               773
                                                                                                                                                                                                                                                                                                                                 Ś
                                                                                                                                                                                                                                                                                                                                                                             22
                                                                                                                                                                                                                                                                                                                                                                                                                          42
                                                                                                                                                                                                                                                                                                                                                                                                                                                                     62
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           593
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  82
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        553
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   713
                                                                                                                                                                                                                                                                           Query Match:
                                                                                                                                                               source
                      DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  LOCUS
DEFINITION
                                                                                                                                                                                   BASE COUNT
                                                                                                                                                                                                                                                                  Best Local
                                                                               ORGANISM
                                                                                                                                                                                                                                 Pred. No.:
                                                                                                                           TITLE
JOURNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            RESULT 2
AF268278
                                                                                                      REFERENCE
                                                                                                                 AUTHORS
                                                                                                                                                  FEATURES
AR179057
                                                                                                                                                                                             ORIGIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                       d
                                                                                                                                                                                                                                                                                                                                                                                                  a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ő
                                                                                                                                                                                                                                                                                                                                  õ
                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                         ŝ
                                                                                                                                                                                                                                                                                                                                                                                                                                              a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   음
```

```
OKNIGEDLHNOIGGWATCVPEDOLLYKGGSIESCKWGGYQFKESEGLPHYPFGKKL
ENETGYRLVDSTSCNREGVALVPQGTLKCKIGKTTVQVIAMDTKLGPWPCRPYEIISS
EGPVEKTACTRYTKTLKNKYPEPRDSYFQQYMLKGEYQYWPLLEVDHHNDYFAESI
LVVVVALLGGTRYVLMILYTWYLSEQKALGIQYGSGEVVAMGHLLTHNNIEVTYFLL
LYLLIREESVKRWALLLHTILVHPIKSYIYILLMIGDYVKADSGGOEYLGKIDLCY
TVVLIVIGLIIARRDPTIVPLVTIMAALRYTELTHQPGVDIAVAWTITLLMVSYYDD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             KNCTPACLPKNTKIVGPGKFDTNAEDGKILHEMGGHLSEVLLLSLVVLSDFAPETASV
MYLILHFSIPQSHVDVMDCDKTQLNLTVELTTADVIPGSVWNLGKWVCIRPNWMPYET
TVYLAFBESGOVKLVLRALRDLTRINNAATTTAFLICLYKTRGWWYGILMLLLT
TVOTLAPEDSGOVKLVLRALRDLTRINNAATTTAFLICLYKTRGWWYGILMLLLIT
TVOTLAFEDSGOVKLDTWVIAMCDGKLL
YLQRCTRETYKIALITHALPTSVVFKKLFPGRKQEDVYEMNDNFEPGLCPCDAKPIV
RGKFNTTLLNGPAFQWVCPIGWTGTVSCTSFNMDTLATTVVRTYRRSKPFPHRQGCIT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             LTLDFMYYMERKVIEEISGGTNIISRLVÄALIELMWSMEEEESKGLKKFYLLSGRLRN
LIIKHKVRNETVASMYGEEEVYGMPKIMTIIKASTLSKSRHCIICTVCEGREWKGGTC
PKCGRHGKPITCGMSLADFEERHYKRIFIREGNFEGMCSRCGGKHRRFEMDREPKSAR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                GARCYVLNPEAVNISGSKGAVVHLOKTGGEFTCYTASGFPAFFDLKNLKGWSGLPIFE
ASSGRVVGRVKVGKNEESKPTKIMSGIQTVSKNTADLIEMVKKITSMNRGDFKQITLA
TGAGKTTELPKAVIEEIGRHKRVLVLIPLRAAAESVYQYMRLKHPSISFNLRIGDMKE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 GKTKSKNTODGLYHNRNRPOESRKLEKALLAMAIIAIVLFQYTMGENITQWNLODNG
TEGIQRAMPORGVNNSLHGIWPEKICTGVPSHLATDIELKTIHGMADASEKTNYTCCR
LORHEWNKHGMCNWYNIEPHILVMRTYGANLTEGOPPRECAYTCRYDRASDLNYYTGA
RDSPTPLTGCKKGKNESFAGILMRGPCNFEIAASDVLFKEHERISMFQDTTLYVYDGI
TNSLEGARGGTAKLTTWLGKOLGILGKKLENKSKTWFGAYAASPYCDVDRKIGYIWYT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       YFRYKWIQCILSLVSGVFLIRSLIYLGRIEMPEVTIPNWRPLTLILLYLISTTYTR
WKVDVAGLLLQCVPILLLVTTLWADFLTLILILPTYELVKLYYLKTVRTDIERSWLGG
IDYTRVDSIYDVDESGEGVYLFPSRQKAQGNFSILLPLIKATLISCVSSRWQLIYMSY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              YCAECNRLHPAEBGDEWAESSMLGLKITYFALMDGKVYDITEWAGCQRVGISPDTHRV
PCHISFGSRWPFRQEYNGFVOYTARGOLFLRNLPVLATKVKMLMVGNLGBEIGNLEHL
GWILRGPAVCKKITEHBKCHINILDKLTAFFGIMPRGTTPRAPVRFPTSLLKVRRGLE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           GDMATGITYASYGYFCQMPQPKLRAANVEYSYIFLDEYHCATPEQLAIIGKIHRFSES
IRVVAMTATPAGSVTTTGQKHPIEEFIAPEVMKGEDLGSQFLDIAGLKIPVDEMKGNM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          LPDLDTVIDTGLKCEKRVRVSSKIPFIVTGLKRMAVTVGEQAQRRGRVGRVKPGRYYR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            SQETATGSKDYHYDLLQAQRYGIEDGINVTKSFREMNYDWSLYEEDSLLITQLEILNN
                                                                                                                                                                                                                                                                                                                                                                                                                                          (bases 1 to 12734)
Lai.V.C.H. and Hong,Z.
Lai.C. Submission
Submitted (16-MAY-2000) Antiviral Therapy, Schering-Plough Research
Institute, 2015 Galloping Hill Road, Kenilworth, NJ 07033-0539, USA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         TGRDKNQYBGEVQIVSTATQIFLATCINGVCWTVYHGAGTRTIASPKGPVIQMYTNVD
QDLVGWPAPQGSRSLTPCTCGSSDLYLVTRHANVIPVRRRGDSRGSLLSPRPISYLKG
SSGGPLLCPAGHAVGLFRAAVCTRGVAKAVDFIPVENLETTTRSGSGADTEDVVCCSM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         TGWAYTHQGGISSVDHVTAGKDLLVCDSMGRTRVVCQSNNRLTDETEYGVKTDSGCPD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     LVFVPTRNMAVEVAKKLKAKGYNSGYYYSGEDPANLRVVTSQSPYVIVATNAI ESGVT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          /translation-"MELNTNEGSGSVVIVGRIVLSGSGSITACAQQTRGLLGCKITSI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SYSDTREEGATKKKTOKPDRLERGKWKIVPKESEKDSKTKPPDATIVVEGVKYQVRK
    Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
Lai, V.C., Zhong, W., Skelton, A., Ingravallo, P., Vassilev, V., Donis, R.O., Hong, Z. and Lau, J.Y. Generation and characterization of a hepatitis C virus NS3 protease-dependent bovine viral diarrhea virus J. Virol. 74 (14), 6339-6347 (2000)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 /organism="Pestivirus type 1"
/mol_type="genomic RNA"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        386. .12508
/codon_start=1
/product="polyprotein"
/protein_id="AAF82566.1"
/db_rref="G1:9049957"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       /mol_type="genomic RN2/db_xref="taxon:11099"
```

```
a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 d
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                            g
                                                                                                    ò
                                                                                                                                    g
VALFGYVGYOALSKRHVPMITDIYTIEDORLEDTTHLQYAPNAIKTDGTETELKELAS
GDVEKIMGAISDYAAGGLEFVKSQAEKIKTAPLFKENAEAAKGYVOKFIDSLIENKEE
IIRYGIMGTHTALKSIAAAKLGHFRAFALTYLKMLAFGGESVSUHVAQAVDLVYYYV
MNEPSFPODETQOETQREYASLFISALATYTYKTNYHNISKVVEPALAYLPYYAY
LAKPSFFDOETQOETQOETGAETATYTYKTNYHNISKVVEPALAYLPYYAY
LAKPTPTRLESVVILSTTYKTYLSIRKGKSDGLLGTGISAAMEILSQNPVSVGISVM
LGYOPLRLIYHLYGVYYKGWENKELSFRTAGNILFTIMFEAFELLGWDSOGKIRNLSGN
YILDLIYGLHKOINGLKKMVLGWAPAPFSCDWTPSDERIRLPTDNYLRVETRCCGY
EMKARYNGGKLKTYVEESGPPLCRNRPGRGPWYRYYYDDNLESK IN PAYLLEGOVE
HYKGYTAKIDYSKGMLLATDKWAPEHGVITRLELVHRNLEBKRIPTATYTWL
AYTEVNEDVATTKVEESGPPLCRNRPGRGPWYRYYYDDNLESK IPPATYTYWL
AYTEVNEDWATTKVEESRYDLENDEVDOETSEWATTGETATYTYWL
AYTEVNEDWATTKRONGCAFTYDLISMLTELELVHRNLEBKRIPTATYTYWL
AYTEVNEDWATTKRONGCAFTYDLISMLTRIELVHRNLEBKRIPTATYTYWL
AYTEVNEDWATTKRONGCAFTYDLISMLTRIELVHRNLEBKRIPTATYTYWL
AYTEVNEDWATTKRONGCAFTYDLISMLTRIELLROPPATTGATGREGKKRR
AKGOPKROYTKRAVRNIIEDQKKONDEIPRWASDPPELENDROGE
VKDOAKALGATDOTRIIKKIRPGGNOTTTTERIHNBARPFIMILGSRNSIS
NRAKTARNINLIKKIRROGGNYPGGOTOTHTLTERIHNBARPFIMILGSRNSIS
KSKNCHMASAYOLAGGNWBEDGGSRYTAMAKISSWFLOASNKOMSLIPFFEELLRCPPAT
KSKNCHMASAYOLAGGNWBEDGGGYHACTIFRRENINKHLEIFTTAOPPTKRYY
VKDTYLIRBHKWHITKKIRROGGNHYRKKLLRPGKLSRCLIPFTETTIAOPPTLNIS
AGIRLEKLPIVRAOTDTRTFHEAIRNINGSTNONNY
VGTYNTENTRYMINGGVLOSSTNONPELHNKLLEIFTTAOPPTLKYTTTA
YGEVTWRODENGENGNYBRYGNYSTNONNY
RENNENNYMINGGVLOSSTNONPELHNKLLEIFTTAOPPTLKYT
YGEVTWRODENGENGNYBRYGNYSTNONNYNGOPVYDEY
                                                                                                                                                                                                                                                                                                                                                                                        TPLFNIFDKVRREMDSFNEPVAVSFDTRAMDTOVTSKDLQLIGEIOKTYTKKEMHKFI
TTDHMTEVPVTTAMCEVYTRAGGNSOPDTSAGNSMLWTTMMYRECESTCOPPK
SFNRVARIHVGENDEFLITEKGLGLKFANKGMOILHEAGKPOKITEGEKMKVAYRFED
IEFCSHTPVPVRWSDNTSSHMAGRDTAVILSKMATRLDSSGERGTTAYEKAVAFSFLL
MYSMPLVRRICLLVLSGOPETDPSKHAYTYTKADFOKNOVIRNEELKRTGFEK
LAMLANLSLSTLGINTHTRKR I 1000VAIGREGNMLVNADRLISSEKRTGFEK
LIQGKHYEQLQLRTFTNPVMGVGTERYKLGPIVNLLRRLKILLMTAVGVSS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               772
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              161
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 652
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    101
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      532
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              81
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       GlySerValValIleValGlyArgIleAsnLeuSerGlyAsp-----ThrAlaTyr
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ProLysGlyProVallleGlnMetTyrThrAsnValAspLysAspLeuValGlyTrpPro
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              AlaProGlnGlySerArgSerLeuThrProCysThrCysGlySerSerAspLeuTyrLeu
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ValThrArgHisAlaAspValIleProValArgArgArgGlyAspSerArgGlySerLeu
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       LeuSerProArgProIleSerTyrLeuLysGlySerSerGlyGlyProLeuLeuCysPro
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  GGTAGTGTTGTTATTGTTGGTAGAATTGTTTTATCTGGTAGTGGTAGTATCACGGCGTGC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    AlaGlnGlnThrArgGlyGluGluGlyCysGlnGluThrSerGlnThrGlyArgAspLys
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CysIleAsnGlyValCysTrpThrValTyrHisGlyAlaGlyThrArgThrIleAlaSer
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1 others
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                12734
178
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Length:
Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          3293 g 2802 t
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (1-12734)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             2608 c
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       US-09-965-594-14 (1-197) x AF268278
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            2.55e-63
910.50
93.85%
91.28%
88.23%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Percent Similarity:
Best Local Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           4030
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            533
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                833
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   22
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     173
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         42
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 593
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    82
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            102
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                713
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    122
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            142
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              62
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Query Match:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ..
80
.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         3'OTR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        BASE COUNT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ORIGIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   _{\text{Pred}}
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Ω
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
```

```
1046
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        PAT 16-MAY-2001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         173
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    94 AspLysAspLeuValGlyTrpProAlaProGlnGlySerArgSerLeuThrProCysThr 113
             AlaGlyHisAlaValGlyIlePheArgAlaAlaValCysThrArgGlyValAlaLysAla 181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                      93
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                154 SerGlyGlyProLeuLeuCysProAlaGlyHisAlaValGlyIlePheArgAlaAlaVal
                                                                                                                                                                                                                                                                                                                                                                                                                                        74 AlaGlyThrArgThrIleAlaSerProLysGlyProValIleGlnMetTyrThrAsnVal
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               114 CysGlySerSerAspLeuTyrLeuValThrArgHisAlaAspValIleProValArgArg
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           134 ArgGlyAspSerArgGlySerLeuLeuSerProArgProlleSerTyrLeuLysGlySer
                                           ValAspPheIleProValGluSerLeuGluThrThrMetArgSer 196
                                                                                                               linear
                                                       5360
175
9
11
11
                                                                                                                                                                                                                                                                                                                        Length:
Matches:
Conservative:
Mismatches:
                                                                                                                                                                                                                                                                                                                                                                                                                          3 LysLysGlySerValVallleValGlyArgIleAsn--
                                                                                                              DNA
                                                                                                             5360 bp us 6150087.
                                                                                                                                                                                                                                                                                1145
                                                                                                                                                                                                                                                                                                                                                                                Gaps:
                                                                                                                                                                                                 1 (bases 1 to 5360)
Chien, D.Y.
NANBV diagnostics and vaccines
Patent: US 6150087-A 53 21-NOV-2/
Location/Qualifiers
1.5360
                                                                                                                                                                                                                                                                                                                                                                                                    (1-5360)
                                                                                                                                                                                                                                                                   /organism="unknown"
1623 c 1532 g
                                                                                                                                                                                                                                                                                 ρ
                                                                                                            AR118686
Seguence 53 from patent
AR118686
AR118686.1 GI:14100596
                                                                                                                                                                                                                                                                                                                                                                                                      US-09-965-594-14 (1-197) x AR118686
                                                                                                                                                                                                                                                                                                                       5.32e-63
901.50
90.20%
85.78%
87.35%
                                                                                                                                                                                        Unclassified
                                                                                                                                                                                                                                                                                                                                                Percent Similarity:
Best Local Similarity:
                                                                                                                                                                                                                                                                               1060 a
                                                                                                                                                                 Unknown
                                                                                                                                                                                                                                                                                                                Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1287
                                           182
162
                                                                                                                                                                                                                                                                                                                                                                      Query Match:
                                                                                                                                 ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
                                                                                                                                                                                                                                                                                                                            ..
Q
                                                                                                                        DEFINITION
                                                                                                                                                                                                                                                                               BASE COUNT
                                                                                                                                                                                                             AUTHORS
TITLE
JOURNAL
FEATURES
                                                                                       RESULT 3
AR118686
                                                                                                                                                                                                  REFERENCE
                                                                                                                                                                                                                                                                                          ORIGIN
                                                                                                                                                                                                                                                                                                                                      Score:
                                                                                                                                                                                                                                                                                                                          Pred
```

	OKCANISM UNKNOWN  URCLASSITIEG.  REFERENCE 1 (bases 1 to 5360)  ACTHORS Houghton, M. Choo., C. and Kuo.G.  JOURNAL Patent: WO 8004669-A 8 01-JUN-1989;  FEATURES Location/Qualifiers  Source 1. 5360  Acrganism unknown 1061 a 1623 c 1533 g 1143 t  ORIGIN	ignment Scores: 5.32e-63 ad. No.: 901.50 forest Similarity: 90.20% st Local Similarity: 85.78% ery Match: 6	US-U9-965-594-14 (1-197) x IU932B (1-536U) Qy 3 LysLysGlySerValValIleValGlyArq1leAsn	867	0y 15LeuSerGlyAspThrAlaTyrAlaGlnGlnThrArgGlyGluGluGlyCysGlnGlu 33	Oy 34 ThreatInthrellyArgAspLysAsnelnValGluClyGluValGlnIleValSerThr 53	Oy 54 AlaAlaGlnThrPheLeuAlaThrCysIleAsnGlyValCysTrpThrValTyrHisGly 73   1111111111111111111111111111111111	Oy 74 AlaGlyThrargThrilealaSerProLysGlyProValileGlnMetTyrThrasnVal 93   110   1111   1	Qy 94 AspLysAspLeuValGlyTrpProAlabroGlnGlySerArgSerLeuThrProCysThr 113	Oy 114 CysGlySerSerAspLeuTyrLeuValThrArgHisAlaAspValIleProValArgArg 133	Oy 134 ArgGlyAspSerArgGlySerLeuLeuSerProArgProlleSerTyrLeuLysGlySer 153	SerGlyGlyProLeuLeuCysProAlaGlyHisAlaValGlyIlePheArgAlaAlaVal 173
Op         1347 TCGGGGGTTCCCTGTTGTGCCCGGGGGCACGCGGGGCATATTTAGGGCCGCGGTG 1406           QY         174 CysThrArgGlyValAlaLysAlaValAspPheIleProValGluSerLeuGluThrThr 193           I         11407 TGCACCGTGGAGTGGCTAAGGCGGTGGACTTTATCCTGTGGAGACTAGAGACACC 1466           QY         194 MetArgSerPro 197           I         111111111111111111111111111111111111	REFERENCE Unclassilited.  AUTHORS Houghton, M., Choo, QL. and Kuo, G.  TITLE Namby diagnostics and vaccines JOURNAL Patent: EP 0318216-A1 48 31-MAY-1989; FRATURES Location/Qualifiers Source /organism="unknown" BASE COUNT 1061 a 1623 c 1533 g 1143 t	Scores: 5.32e-63 100.20% 1 Similarity: 90.20% 1 Similarity: 85.78% 6h: 6	US-09-965-594-14 (1-197) x 106434 (1-5360) Qy 3 LysLysGlySerValValIleValGlyArqIleAsn	867	Qy 15LeuSerGlyAspThrAlaTyrAlaGlnGlnThrArgGlyGluGluGlyCysGlnGlu 33	Oy 34 ThrSerGlnThrGlyArgAspLysAsnGlnValGluGlyGluValGlnIleValSerThr 53	Oy 54 AlaAlaGlnThrPheLeuAlaThrCysIleAsnGlyValCysTrpThrValTyrHisGly 73 [11111111111111111111111111111111111	Oy 74 AlaGlyThrArgThrILeAlaSerProLysGlyProVallIeGlnMetTyrThrAsnVal 93 	Qy 94 AspLysAspLeuValGlyTrpProAlabroGlnGlySerArgSerLeuThrProCysThr 113	<pre>Qy 114 CysGlySerSerAspLeuTyrLeuValThrArgHisAlaAspValIleProValArgArg 133</pre>	Qy 134 ArgGlyAspSerArgGlySerLeuLeuSerProArgPro11eSerTyrLeuLySGlySer 153	SerGlyGlyProLeuLeuCysProAlaGlyHisAlaValGlyIlePheArgAlaAlaVal

us-09-965-594-14.rge

0y   154   SerGlyGlyProLeuLeuCysProAlaGlyHisAlavalGlyIlePheArgAlaAlaval 173	194 MetargSerPro 197 	RESULT 7 106440 LOCGS LO	: ty: arity: 6	US-09-965-594-14 (1-197) x 106440 (1-6785)  Qy	Db 1140 CGCAGGGCCGGGAGATACTGCTCGGGCCGGTGGAATGGTCTCCAAGGGTGGAAGG 1199 Oy 15LeuSerGlyAspThrAlaTyrAlaGlnGlnThrArgGlyGluGluGlyCysGlnGlu 33   1	34 ThrSerGlnThrGlyArgAspLySAsnGlnValGluGlyGluValGlnIleValSerThr 53	54 AlaalaGlnThrPheLeuAlaThrCysIleAsnGlyValCysTrpThrValTyrHisGly	/4 AladyThrangThrileAlaSerProLysGlyProValileGlmMeTlyThrashVal 93 [11]	DD 1440 GACCAMGACCTIGIGGGCTGGCCCCCCCCCAMGGTAGCCGCTCATIGACACCTTGCACT 1499  Qy 114 CysGlySerSerAspLeuTyrLeuValThrArgHisAlaAspValIleProValArgArg 133	Qy 134 ArgGlyAspSerArgGlySerLeuLeuSerProArgProlleSerTyrLeuLysGlySer 153 
11111111111111111111111111111111111111	Oy 194 MetArgSerPro 197		Automent Scores: 6.86e-63 Length: 6785 Score: 901.50 Matches: 175 Percent Similarity: 90.20% Conservative: 9 Best Local Similarity: 85.78% Mismatches: 11 Query Match: 87.35% Indels: 1 DB: 6.735% Gaps: 1 US-09-965-594-14 (1-197) x AR118692 (1-6785)		15LeuSerGlyaspThrAlaTyrAlaGlnGlnThrArgGlyGluGluGlyCysGlnGlu 33	34 ThrSerGlnThrGlyArgAspLysAsnGlnValGluGlyGluValGlnIleValSerThr 5   11111   111	Oy SA ALAALAGADTINFPRE-BENALATING STLEASENCLYVALCYSTEPTINFVALTYTHISGLY 73  1320 GCTGCCCAAACCTCCTGGCAACGTGGATCAATGGGGTGTGCTGGACTCTACCACGG 1379  Oy 74 AlaGlyThIAAGTTILEALBSEPFOLYSGLYPPOVALILGGINMETTYTHYASNVAL 93	Db 1380 GCCGGAACGAGGACCTCGCCAAGGGTCCTCATCTTTTTTTT	114 CysGlySerSerAspLeuTyrLeuValThrArgHisAlaAspVallleProValArgArg 13	Oy 134 ArgolyAapSerArgolySerLeuLeuSerProArgProI1eSerTyrLeuLySolySer 153 

٥y	15	54 S	
q	163	1620 TCGGGGGGTCCGCTGTTGTGTCCCCCGCGGGCACGCCGTGGGCATATTAGGCCCGCGGT	
Οy	17	174 CysthrargGlyvalAlaLussAlavalAspPheIleProvalGluSe	rLeuGluThrThr 193
qq	168	1680 TGCACCGTGGAGTGGCTAAGGCGGTGGACTTTATCCCTGTGGAGAACCTAGAGAACC	CCTAGAGACAACC 1739
\$ 6	15	194 MetArgSerPro 197 	
3	_	C/1 933331000010 01/	
RESUJ 1064- 100CU 10CU DEFIJ ACCES VERS SOURC SOURC	S NITIC SSION SSION DRDS	7 106440 6785 bp DNA linear IO Sequence 54 from Patent EP 0318216. ON 106440 GI:590312 S UNKNOWN.	PAT 02-DEC-1994
REFE AU TI JOI FEATI	RENCI TLE TLE URNAI GRES	Unclassified.  U (bases 1 to 6785) Houghton, M., Choo, Q. Nanby diagnostics ar Patent: EP 0318216 ** Location/O.	
BASE ORIG	COO	/organismunknown 1392 a 2050 c 1914 g	
Align Pred Score Perce Best Query DB:	nment S No.: e: ent Sim Local	nt O.: Síu cal	
-SD	5-596-60	- 59	
δý		3 LysLysGlySerValValIleValGlyArglleAsn	
QC	114	140 CGCAGGGCCGGGAGATACTGCTCGGGCCAGCCGATGGAATGGTCT	CCAAGGGGTGGAGG 1199
Oy Db	120	15LeuSerGlyAspThrAlaTyrAlaGlnGlnThrArgGly 111:::	GluGluGlyCysGlnGlu 33          CTCCTAGGGTGCATAATC 1259
Qy		34 ThrSerGlnThrGlyArgAspLysAsnGlnValGluGlyGluValGlnIleValSerThr	nilevalSerThr 53
qq	12(	260 ACCAGCCIAACIGGCCGG	GATTGTGTCAACT 1319
oy		ry (	y 73
a a	13.	320 GC	
QY Db	136		tTyrThrAsnval 93 
٥y		94 AspLysAspLeuValGlyTrpProAlaProGlnGlySerArgSerLe	uThrProCysThr 113
qo	14	1440 GACCAAGACCTTGTGGGCTGGCCCGCTCGGAAGGTAGCCGCTCATTGACACCTG	GACACCCTGCACT 1499
0y	1	114 CysGlySerSerAspLeuTyrLeuValThrArgHisAlaAspValIlePr	eProvalArgArg 133
qq	15(	500 rececrecre	cerececce
δλ	Ħ	134 ArgGlyAspSerArgGlySerLeuLeuSerProArgProlleSerTyrLeuLysGlySer	rLeuLysGlySer 153
q	15(	260	CTIGAAAGGCICC 1619

154 ; 1620 ' 1620 ' 1680 ' 1740 ; 174	Oy 34 Threarg.  Db 1785 ACCAGCC Oy 54 AlaalaG Oy 74 AlaGlyT) Db 1845 GCTGCCC Oy 74 AlaGlyT) Db 1905 GCGGAAA Oy 94 ABLySA Oy 94 ABLySA Oy 114 CYSGLySA Oy 114 CYSGLySA Oy 114 CYSGLYSA Oy 134 ArgGlySA
154   SerGlyGlyProLeuLcuCySProAlaclyHisAlavalGlyIlePheArgAlaAlaval   173     1620   IGGGGGGTCGGTGTTGTCGCCGGGGGGCAGCGCTGGGCATALTAGGGCCGCGGTG   1679     174   CysThrangGlyValAlaLysalavalAspheileProValGluSerLeuGluThThr	11

```
:::
sccccatcacggcgtacgcccagcagacaagggcctcctagggtgcataatc 1784
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 AspSerArgGlySerLeuLeuSerProArgPro1leSerTyrLeuLysGlySer 153
                                                                                                                          PAT 16-MAY-2001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ThrArgThrIleAlaSerProLysGlyProValIleGlnMetTyrThrAsnVal 93
                                                                                                                                                                                                                                                                                                                                                                                                                SerGlyAspThrAlaTyrAlaGlnGlnThrArgGlyGluGluGlyCysGlnGlu 33
                                                                                                                                                                                                                                                                                                                                                                            31ySerValValIleValGlyArgIleAsn-----
                                                                                                                          linear
                                                                                                                                                                                                                                                                                             7310
175
9
11
11
                                                                                                                                                                                                                                                                                             Length:
Matches:
Conservative:
Mismatches:
Indels:
Gaps:
                                                                                                                           DNA
                                                                                                                                                                                                                agnostics and vaccines
US 6150087-A 74 21-NOV-2000;
Location/Qualifiers
1..7310
                                                                                                                         7310 bp 1
74 from patent US 6150087.
                                                                                                                                                                                                                                                           1539
                                                                                                                                                                                                                                                                                                                                                           197) x AR118696 (1-7310)
                                                                                                                                                                                                                                                 /organism="unknown"
a 2220 c 2056 g
                                                                                                                                                   1 GI:14100606
                                                                                                                                                                                                                                                                                             7.44e-63
901.50
90.20%
85.78%
87.35%
                                                                                                                                                                                        fied.
s 1 to 7310)
                                                                                 SerPro 197
```

Db   2085 CGGGGTGATAGCAGGGGCAGCCTGCTGTG   154 SerClyGlyProLeuLeuCysProAlad   154 SerClyGlyProLeuLeuCysProAlad   154   155	RESULT 11 HPCPOLYP LOCUS DETINITION HPCPOLYP ACCESSION M32084.1 GI:329875 VERMORDS SOURCE COURS HPATILIS C VITUS OUGHAILS C VITUS OURCANISM HPATILIS C VITUS OURCANISM HPATILIS C VITUS OURCANISM HPATILIS C VITUS VITUSES; SSRNA POSITIVE-STRAND HEFERENCE 1 (bases 1 to 7310) AUTHORS CHOO, -L., RICHMAN, K. and Han HEFERENCE 1 (bases 1 to 7310) AUTHORS CHOO, -L., RICHMAN, K. and Han AUTHORS CHOO, -L., RICHMAN, R. and Han AUTHORS OUNDALISHED (1990) COMMENT R9-1 through 15e, isolated from Plasma.	Draft entry and printed sequen M.Houghton, 22-FEB-1990. Chiro Emeryville CA 94608. Location/Qualifiers Location/Qualifiers Jorganism-"Hepatitis (/mol.type-"genomic RN //mol.type-"genomic RN //mol.type-"geno	/ COUNTY OF THE PROPERTY / DEVICE IN 1998   F. / LT STATE   F.   STATE   STATE   F.   STATE   STATE	GAVYTFYGWWPLLLLLLALPO LWMLQYFLTRVEAGLHWWIPP ILOASLLKVPYFWRVOGLIRE WAHNGLRDLAVAVEPVVFSOM GWVSKGRRILLAPTTAYAOQTR NOVCWYVYRGAGTRIASPKG VTRHADVI PVRRRODSRGSLL KAVDFIPVENLETTWRSEVFT KAVDFIPVENLETTWRSEVFT KAVDFIPVENLETTWRSEVFT GGAYDI ILOEGGISTANGST	VALSTTGEIPFYGKAIPLEVI VSVIPTSGDVVVATDALHTG AVSRTQRRGRTGRGKGITFRE LRAYMNTFGLPVCQDHLEFWE AQAPPFSWDQMRKCLIRLEFW VYTSTWTLVGGVLALAAYCL QHLPYIEGGMALAEQFROKAL GIQYLAGLSTLFGNPAIASLM FYGGALAVAGGVGLKKTLI LSPGAIVVGVVCAA GIRRHYG TAILSSLTVTQLLRRHOWIS LPGIPFVSCQRGYKGVWRVDG
Db 2085 GGGGGGGAAAGCAGGGCGGGCCCGGGCCCGATTCCTACTTGAAAGGCTCC 2144  Qy 154 SerGlyGlyProLeuLcuCySProAlaGlyHisAlaValGlyIlePhcArgAlaVal 173	Db   2265 ATGAGGTCCCG 2276   RESULT 10   19931   7310 bp   DNA   linear PAT 02-DEC-1994   LOCUS   109331   7310 bp   DNA   linear PAT 02-DEC-1994   DEFINITION   Sequence 15 from Patent WO 8904669.   ACCESSION   109331   G1:587966   KETWORDS   109331   G1:587966   KETWORDS   Unknown   109331   G1:587966   C1   G1:587966   G1:5879666   G1:587966   G1:5	Alignment Scores: 7.44e-63 Length: 7310  Pred. No.: 901.50 Matches: 175  Score: 90.20% Conservative: 9  Best Local Similarity: 85.78% Mismatches: 1  Ouery Match: 6  US-09-965-594-14 (1-197) x 109331 (1-7310)	LysLysGlySerValValIleValGlyArgIleAsn	Oy         34 ThrSerGlnThrGlyArgAspLy9AsnGlnValGluGlyGluValGlnIleValSerThr 53           IIIIII         IIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIII	1905   GCGGGAACGAGGACCATCGGGTCACCCAAGGGTCTGTCATCCAGATGTATACCAATGTA   1964     94   AsplysaspleuValG1YTrpProalaProGlnG1ySerArgSerLeuThrProCysThr   113

```
q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Qγ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                οχ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    οý
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
          PEFFEELDOYRELHERANGERENCYBERGYSTRYCHIETPYGGOLPCEPEPDYAUTSHITD DRIVENCYBERGYBERGYBERGYBY TO THE TOTAL TABLAGGRELANGER TO THE TOTAL TABLAGGRELANGER TO THE TOTAL TABLAGGRELANGER TO THE TOTAL THE OPERIT DESTILET THE OPENITARY TO THE TOTAL THE SEGSST SCHITGHTTY TO THE TOTAL THE SEGSST SCHITGHTY TO THE TOTAL THE TOTAL THE THE THE TOTAL THE SEGSST SCHITGHTY TO THE TOTAL 
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1904
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               2084
PINAYTTGPCTPLPAPNYTFALWRVSAEEYVEIRQVGDFHYVTGMTTDNLKCPCOVPS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ||||::
|TGCTGGCGCCCATCACGGCGTACGCCCAGCAGACAAGGGGCCTCCTAGGGTGCATAATC ||194
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 AlaGlyThrArgThrIleAlaSerProLySGlyProValIleGInMetTyrThrAsnVal 93
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1665 CGCAGGGCCGGGAGATACTGCTCGGGCCAGCCGATGGAATGGTCTCCAAGGGGTGGAGG 1724
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           113
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         133
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              53
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            73
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ---LeuSerGlyAspThrAlaTyrAlaGlnGlnThrArgGlyGluGluGlyCysGlnGlu 33
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     AlaAlaGlnThrPheLeuAlaThrCysIleAsnGlyValCysTrpThrValTyrHisGly
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               154 SerGlyGlyProLeuLeuCysProAlaGlyHisAlaValGlyIlePheArgAlaAlaVal
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        134 ArgGlyAspSerArgGlySerLeuLeuSerProArgProlleSerTyrLeuLysGlySer
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ThrSerGlnThrGlyArgAspLysAsnGlnValGluGlyGluValGlnIleValSerThr
                                                                                                                                                                                                                                                                                                                                                                                  7310
175
9
11
11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    LysLysGlySerValValIleValGlyArgIleAsn----
                                                                                                                                                                                                                                                                                                                                                                                Length:
Matches:
Conservative:
Mismatches:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           US-09-965-594-14 (1-197) x HPCPOLYP (1-7310)
                                                                                                                                                                                                                                                                                                        2058 g
                                                                                                                                                                                                                                                                                   SIEPLDLPPIIQRL
2218 c 20
                                                                                                                                                                                                                                                                                                                                                                                7.44e-63
901.50
90.20%
85.78%
87.35%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            11111111111
2265 ATGAGGTCCCG 2276
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       MetArgSerPro 197
                                                                                                                                                                                                                                                                                                                                                                                                               Percent Similarity:
Best Local Similarity:
Query Match:
DB:
                                                                                                                                                                                                                                                                                                        ø
                                                                                                                                                                                                                                                                                                      1495
                                                                                                                                                                                                                                                                                                                                                                 Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              15
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1725
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            34
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1785
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          54
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1845
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           74
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1905
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1965
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   194
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       46
                                                                                                                                                                                                                                                                                                    BASE COUNT
ORIGIN
                                                                                                                                                                                                                                                                                                                                                                                    Pred. No.:
                                                                                                                                                                                                                                                                                                                                                                                                       Score:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
```

```
GACCAAGACCTTGTGGGCTGGCCCGCTCCGCAAGGTAGCCGCTCATTGACACCCCTGCACT 3030
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     TCGGGGGGTCCGCTGTTGTGCCCCGCGGGGCACGCGTGGGCCATATTTAGGGCCGCGCGG 3210
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        113
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                173
              PAT 16-MAY-2003
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       133
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   153
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CysThrArgGlyValAlaLysAlaValAspPheIleProValGluSerLeuGluThrThr 193
                                                                                                                                                                                                                                                                                                                                                                                                                       53
                                                                                                                                                                                                                                                                                                                                                                                                                                                                   73
                                                                                                                                                                                                                                                                                                                                                                           33
                                                                                                                                                                                                                                                                                                                                                                                                                                                                AspLysAspLeuValGlyTrpProAlaProGlnGlySerArgSerLeuThrProCysThr
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SerGlyGlyProLeuLeuCysProAlaGlyHisAlaValGlyIlePheArgAlaAlaVal
                                                                                                                                                                                                                                                                                                                                                                          ---LeuSerGlyAspThrAlaTyrAlaGlnGlnThrArgGlyGluGluGlyCysGlnGlu
                                                                                                                                                                                                                                                                                                                                                                                                                     ThrSerGlnThrGlyArgAspLysAsnGlnValGluGlyGluValGlnIleValSerThr
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CysGlySerSerAspLeuTyrLeuValThrArgHisAlaAspValIleProValArgArg
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ArgGlyAspSerArgGlySerLeuLeuSerProArgProlleSerTyrLeuLysGlySer
              linear
                                                                                                                                                                                                                           8316
175
9
11
                                                                                                                                                                                                                           Length:
Matches:
Conservative:
Mismatches:
Indels:
Gaps:
              ONA
                                                                                                                                                                                                                                                                                                                               LysLysGlySerValVallleValGlyArgIleAsn
                                                                                                 Chien, D.Y.

NANBY diagnostics and vaccines
Patent: US 6150087-A 88 21-NOV-2000;
Location/Qualifiers
1. 8316
           8316 bp
US 6150087.
                                                                                                                                                                                                                                                                                                        (1-8316)
                                                                                                                                                                    /organism="unknown"
2529 c 2345 g
                                                                                                                                                                                  φ
                       88 from patent
                                                                                                                                                                                                                                                                                                        US-09-965-594-14 (1-197) x AR118703
                                             AR118703.1 GI:14100613
                                                                                                                                                                                                                          8.55e-63
901.50
90.20%
85.78%
87.35%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ATGAGGTCCCCG 3282
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        MetArgSerPro 197
                                                                                          Unclassified
                                                                                                                                                                                                                                                  Percent Similarity:
Best Local Similarity:
          AR118703
Sequence 8
AR118703
                                                                                                                                                                                1671 a
                                                                   Unknown
                                                                             Unknown
                                                                                                                                                                                                                   Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                               2731
                                                                                                                                                                                                                                                                                                                                                                                                                                           2791
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        2851
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                2971
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            3031
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         3091
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              3271
                                                                                                                                                                                                                                                                                                                              m
                                                                                                                                                                                                                                                                                                                                                                        15
                                                                                                                                                                                                                                                                                                                                                                                                                     34
                                                                                                                                                                                                                                                                                                                                                                                                                                                                 54
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          46
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     114
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   134
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              154
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     3151
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           174
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 3211
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       194
                                                                                                                                                                                                                                                                        Query Match:
                     DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
                                                                                                                                                           source
                                                                                                                                                                                                                               ..
0N
                                                                                                                         TITLE
JOURNAL
FEATURES
                                                                                                                                                                                BASE COUNT
                                                                                                   REFERENCE
AUTHORS
AR118703
                                                                                                                                                                                           ORIGIN
```

16-MAY-2001

σ

```
54 AlaAlaGlnThrPheLeuAlaThrCysIleAsnGlyValCysTrpThrValTyrHisGly 73
                                                                                                                                                                                                                                                                                                 3332 CGCAGGGGCCGGGAGATACTGCTCGGGCCAGCCGATGGAATGGTCTCCAAGGGGTGGAGG
                                                                                                                                                                                                                                                                                                                                                                                                SerGlyGlyProLeuLeuCysProAlaGlyHisAlaValGlyIlePheArgAlaAlaVal
                             PAT
                              linear
                                                                                                                                                                                                                                                                                LysLysGlySerValVallleValGlyArgIleAsn-----
                                                                                                                                                                                                Length:
Matches:
Conservative:
Mismatches:
Indels:
                                                                                                1 (bases 1 to 9185)
Chien, D.Y.
NANBV disprostics and vaccines
Patent: US 6150087-A 122 21-NOV-2000;
Location/Qualifiers
1. 9185
                         9185 bp D Sequence 122 from patent US 6150087. ARI18722 ARI18722.1 GI:14100632
                                                                                                                                                              1938
                                                                                                                                                                                                                                             Gaps:
                                                                                                                                                                                                                                                             US-09-965-594-14 (1-197) x AR118722 (1-9185)
                                                                                                                                                   /organism="unknown"
2790 c 2608 q
                                                                                                                                                                                               9.52e-63
901.50
90.20%
85.78%
87.35%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           MetArgSerPro 197
                                                                                         Unclassified
                                                                                                                                                                                                                 Percent Similarity:
Best Local Similarity:
                                                                                                                                                              1849 a
                                                                       Unknown
                                                                                                                                                                                        Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    154
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           194
                                                                                                                                                                                                                                                                               m
                                                                                                                                                                                                                                                                                                                                                                                                                                                                             3632
                                                                                                                                                                                                                                   Query Match:
                                                                                                                                              source
                                   DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
                                                                                                                                                             BASE COUNT
ORIGIN
                                                                                                         AUTHORS
TITLE
JOURNAL
                                                                                                                                                                                                  Pred. No.:
          RESULT 14
AR118722
LOCUS
                                                                                                 REFERENCE
                                                                                                                                    FEATURES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                 a
                                                                                                                                                                                                                                                                                                                                 a
                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                     οχ
                                                                                                                                                                                                                                                                                                                                                                                                      셤
                                                                                                                                                                                                                                                                                                                                                                                                                        ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                       qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               οğ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ^{0}
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         셤
                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                   3372
                                                                                                                                                                                                                                                                                       3072
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   3193 GCTGCCCAAACCTTCCTGGCAACGTGCATCAATGGGGTGTGCTGGACTGTCTACCACGGG 3252
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         3433 CGGGGTGATAGCAGGGGCAGCTGCTGTCGCCCGGCCCATTTCCTACTTGAAAGGCTCC 3492
                                                                                                                                                                                                                                                                                                                                                                                                                                                 113
                   PAT 16-MAY-2001
                                                                                                                                                                                                                                                                                                                                                                              73
                                                                                                                                                                                                                                                                       7
                                                                                                                                                                                                                                                                                                                                                                                                                                                 3013 CGCAGGGGCCGGGAGATACTGCTCGGGCCAGCCGATGGAATGGTCTCCAAGGGGTGGAGG
                                                                                                                                                                                                                                                                                                                                                                             54 AlaAlaGlnThrPheLeuAlaThrCysIleAsnGlyValCysTrpThrValTyrHisGly
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       134 ArgGlyAspSerArgGlySerLeuLeuSerProArgProIleSerTyrLeuLysGlySer
                    linear
                                                                                                                                                                                                                                                                     3 LysLysGlySerValValIleValGlyArgIleAsn----
                                                                                                                                                                                                         Conservative:
Mismatches:
Indels:
                                                                                      1 (bases 1 to 8987)
Chien, D.Y.
NANDY diagnostics and vaccines
Patent: US 6150087-A 137 21-NOV-2000;
Location/Qualifiers
1. 8987
                                                                                                                                                                                       Length:
Matches:
                           Sequence 137 from patent US 6150087.
AR118728
AR118728.1 GI:14100638
                                                                                                                                                                                                                                                    US-09-965-594-14 (1-197) x AR118728 (1-8987)
                                                                                                                                        /organism-"unknown"
2735 c 2547 a
                                                                                                                                                                                    9.3e-63
901.50
90.20%
85.78%
87.35%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           3613 ATGAGGTCCCCG 3624
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  MetArgSerPro 197
                                                                                 Jnclassified
                                                                                                                                                                                                    Percent Similarity:
Best Local Similarity:
Query Match:
                                                                                                                                                      N
                   AR118728
                                                                                                                                                     1807
                                                               Unknown
                                                                       Unknown
                                                                                                                                                                               Alignment Scores:
Pred. No.:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 194
                         DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
                                                                                                                                    source
                                                                      ORGANISM
                                                                                                AUTHORS
TITLE
JOURNAL
FEATURES
                                                                                                                                                     BASE COUNT
RESULT 13
AR118728
LOCUS
                                                                                        REFERENCE
                                                                                                                                                             ORIGIN
                                                                                                                                                                                                 Score:
                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ολ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               윱
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 õ
                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                        a
                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      δ
                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                          à
                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                             å
```

```
셤
                                                                                                                                                                                                                                                                                              PAT 16-MAY-2001
                                                                                                                                                                                                                                                                                     34 ThrSerGlnThrGlyArgAspLysAsnGLnValGluGlyGluValGlnIleValSerThr
                              linear
                                                                                                                                                              Length:
Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                               3 LysLysGlySerValVallleValGlyArglleAsn---
                              DNA
                                                                             Unclassified.
1 (bases 1 to 9185)
Chieb. D.Y.
NANBV diagnostics and vaccines
Patent: US 6150087-A 123 21-NOV-2000;
1. 9185
                            AR118723 9185 bp C
Sequence 123 from patent US 6150087.
AR118723.1 GI:14100633
                                                                                                                                      1849
                                                                                                                                                                                                    Gaps:
                                                                                                                                                                                                                 US-09-965-594-14 (1-197) x AR118723 (1-9185)
                                                                                                                              /organism="unknown"
2608 c 2790 g
                                                                                                                                                                9.52e-63
901.50
90.20%
85.78%
87.35%
3932 ATGAGGTCCCCG 3943
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        MetArgSerPro 197
                                                                                                                                                                           Percent Similarity:
Best Local Similarity:
Query Match:
DB:
                                                           Unknown.
                                                                       Unknown
                                                                                                                                                          Alignment Scores:
Pred. No.:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       194
              RESULT 15
AR118723/c
LOCUS
DEFINITION
ACCESSION
KEYMORDS
SOURCE
                                                                       ORGANISM
                                                                                   REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
                                                                                                                                     BASE COUNT
ORIGIN
a
                                                                                                                                                                                                                               ö
                                                                                                                                                                                                                                           a
                                                                                                                                                                                                                                                                        셤
                                                                                                                                                                                                                                                                                                   셤
                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                     q
                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                   d
                                                                                                                                                                                                                                                                                                                                                                                                                                   ð
                                                                                                                                                                                                                                                                                                                                                                                                                                              qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                              õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                           d
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        a
```

ö

õ

```
Search completed: August 31, 2003, 00:46:04
Job time : 2584.57 secs
5254 ATGAGGTCCCCG 5243
```

```
GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.
```

```
(without alignments)
2906.924 Million cell updates/sec
                                                       August 30, 2003, 19:13:57; Search time 182.939 Seconds
OM protein - nucleic search, using frame_plus_p2n model
                                                           Run on:
```

1 MKKKGSVVIVGRINLSGDTA......VAKAVDFIPVESLETTMRSP 197 0.5 7.0 7.0 Xgapop 10.0 , Xgapext Ygapop 10.0 , Ygapext Fgapop 6.0 , Fgapext Delop 6.0 , Delext US-09-965-594-14 1032 **BLOSUM62** Perfect score: Sequence: Scoring table:

5105512 Total number of hits satisfying chosen parameters:

2552756 seqs, 1349719017 residues

Searched:

Minimum DB seq length: 0 Maximum DB seq length: 200000000

Post-processing: Minimum Match 0% Maximum Match 100% Listing first 45 summaries

COMMAND line parameters:

-MODEL-frame+p2n.model -DEV-xlp
-O-/Compand line parameters:
-MODEL-frame+p2n.model -DEV-xlp
-O-/Compand line parametery.fasta\_1.2872
-O-/Compand\_1/GPF0\_app\_001/0809965594/runat\_29082003\_151918\_28302/app\_query.fasta\_1.2872
-DB=N\_Geneseq\_19Jun03 -OFMT-fastap -SUFFIX-ring -MINMATCH=0.1 -LOOPCL=0.0
-LOOPEXT-0 -UNITS-bits -START-1 -END--1 -HR\_MXX=100 -THR\_MIN-0 -ALIGN-15
-LIST-45 -DOCALIGN-200 -THR\_SCORE-pct -THR\_MXX=100 -THR\_MIN-0 -ALIGN-15
-MODE-LOCAL -OUTFNT-pto -NORM-ext -HEAPSIZE-500 -MINLEN-0 -NAXLEN-2000000000
-USER-USSO9965594\_GCGN\_1\_1-1412\_errnat\_29082003\_151918\_28302 -NCPU-6 -ICPU-3
-NO\_MMAP -LARGEQUERY -NEC\_SCORES-0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV\_IRECOTT-120 -WARN\_TIMEOOT-30 -THREADS-1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELEXT=7

/SIDS1/gcgdata/geneseg/genesegn-emb1/NA2000.DAT:\*/SIDS1/gcgdata/geneseg/genesegn-emb1/NA2001A.DAT:\*/SIDS1/gcgdata/geneseg/genesegn-emb1/NA2001B.DAT:\* /SIDSI/gcgdata/geneseq/geneseqn-embl/NA1989.DAT:\*/SIDSI/gcgdata/geneseq/geneseqn-embl/NA1990.DAT:\*/SIDSI/gcgdata/geneseq/geneseqn-embl/NA1991.DAT:\*/SIDSI/gcgdata/geneseq/geneseqn-embl/NA1992.DAT:\*/SIDSI/gcgdata/geneseq/geneseqn-embl/NA1992.DAT:\*/SIDSI/gcgdata/geneseq/geneseqn-embl/NA1993.DAT:\*/SIDSI/gcgdata/geneseqn-embl/NA1994.DAT:\*/SIDSI/gcgdata/geneseqn-embl/NA1994.DAT:\*/SIDSI/gcgdata/geneseqn-embl/NA1995.DAT:\*/SIDSI/gcgdata/geneseqn-embl/N /SIDS1/gogdata/geneseg/genesegn-embl/NA1996.DAT:\*/SIDS1/gogdata/geneseg/genesegn-embl/NA1997.DAT:\*/SIDS1/gogdata/geneseg/genesegn-embl/NA1998.DAT:\* /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1999.DAT:\* 15: 16: 17: 18: 19: 20: 22: :01 Database :

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed,

/SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:\*/SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2003.DAT:\*

and is derived by analysis of the total score distribution.

	יינה ביינה יינה יינה יינה יינה יינה יינה	DE Hepatitis C virus NS4A-NS3 fusion protease coding sequence #3.	TT 19-DEC-2000 (first entry)	.c. AAA73330; x	.AA73330 D AAA73330 standard; DNA; 594 BP.	ESULT 1	ALIGNMENTS		5 692.5 86.5 9365 24 AAD25518 Hepatitis C vi	4 894.5 86.7 1998 20 AAX80355 HCV NS4A-NS3 C	3 895 86.7 2061 24 AAD31767 Hepatitis C v	1 090:3 80:9 10022 21 AAC35212 NUCLEOTIDE S 2 895 86.7 2061 24 AAD34500 Hebatitis C	0 896.5 86.9 12980 24 ABK87286 Hepatitis C	9 896.5 86.9 12980 19 AAV59364 Hepatitis C v	8 896.5 86.9 9646 24 ABK87285 CDNA encoding	6 897 86.9 1933 20 AAX23258 HCV NS3 DNA.	5 897.5 87.0 8316 11 AAQ05955 Hepatitis C	4 898.5 87.1 9185 20 AAX00459 Hepatitis C	2 839 87.1 8145 20 AAXZSZ59 Plasmid pET- 3 898.5 87.1 9185 20 AAXZ6737 Nucleotide s	1 899 87.1 2058 25 ABX14410 DNA encoding HC	0 899 87.1 2058 24 ABK15344 Hepatitis C	8 89 87.1 2058 24 AAD29795 HCV-1 NS3/4	7 900.5 87.3 9502 15 AAQ74770 Hepatitis C 8 899 87.1 549 21 AAA70344	6 901.5 87.4 9401 24 AAD35043 Repatitis C viru	5 901.5 87.4 9401 19 AAV09989 HCV polyprotein	4 901.5 87.4 9401 18 AAT99981 HCV boly	2 901.5 87.4 9400 13 AAQ21744 Compiled	1 901.5 87.4 9185 21 AAA75297 Sense st	0 901.5 87.4 9185 12 AAQ10566 Hepatiti	9 901.5 87.4 9185 11 AA005956 Sepse St	/ 901.5 87.4 8310 21 AAA/5296 CDNA Sequence 8 901.5 87.4 9133 20 AAA/5256 Nuclearida co	6 901.5 87.4 7310 16 AAO98221 Hepatiti	5 901.5 87.4 7310 10 AAN90336 Composit	4 901.5 87.4 7310 10 AAN92106 Combined	3 901.5 87.4 6905 10 AAN92103 Combined	2 901.5 87.4 5360 10 AAN90327 Hepatitis C viru	1 901.5 87.4 5300 10 AAN92097 Combined open	0 921.5 89.3 612 25 ABX15706 Anti-viral synth	930.5 90.2 12734 24 ABA95615 Chimeric Runs	959 92.9 588 21 AAA73328 Henatitis C	963 93.3 594 21 AAA73334 Henatitis C	973 94.3 594 21 AAA7333 Heparitis C	985 95.4 594 21 AAA7332 Hebatitis	993 96.2 594 21 AAA73335 Heparifis C	998 96.7 588 21 AAA73329 Hepatitis C	1015 98.4 594 21 AAA73331 Henatitis C	1 1032 100 0 594 01 ABA72220	t . Score Match Length DB ID	
No.   Score   March Length DB   ID	Ouery  Score March Length DB ID  1012 100.0 594 21 AAA73330  1022 100.0 594 21 AAA73331  933 96.2 594 21 AAA73333  94.3 594 21 AAA73333  95.2 595 21 AAA73334  95.2 595 21 AAA73334  95.2 595 21 AAA73334  95.2 596 10 AAA73344  95.2 597 4 596 10 AAA73249  95.2 597 4 596 11 AAA73344  95.2 598 57 1 508 54 AAA73249  96.2 598 57 1 508 54 AAA73249  96.2 598 57 1 508 54 AAA73349  96.2 598 57 1 508 54 AAA73349  97.2 598 57 1 508 54 AAA73349  98.5 587 1 508 54 AAA73349  99.5 587 1 508 54 AAA733450  99.6 588 569 1 508 54 AAA73340  99.7 588 54 AAA73340  99.7 588 54 AAA73340  99.8 589 580 1 508 54 AAA73340  99.8 589 580 1 508 54 AAA73340  99.8 580 580 580 580 580 580 580 580 580 58	Score Match Length DB ID  1032 100 0 594 21 AAA7333 Hepatiii 1988 4 594 21 AAA7333 Hepatiii 998 96.7 594 21 AAA7333 Hepatiii 909 96.7 594 21 AAA7333 Hepatiii 900 96.7 594 21 AAA7333 Hepatiii 900 96.7 594 21 AAA7333 Hepatiii 901 96.7 594 21 AAA7333 Hepatiii 902 96.7 594 21 AAA7333 Hepatiii 903 96.7 594 21 AAA7333 Hepatiii 903 96.7 594 21 AAA7333 Hepatiii 903 96.7 594 21 AAA7333 Hepatiii 901.5 87.4 5100 AAN9300 Combined 901.5 87.4 5100 AAN9300 Combined 901.5 87.4 5100 ID AAN9300 Combined 901.5 87.1 2058 24 AAD3200 Combined 902.5 87.1 2000 ID AAN9300 Combined 903.8 87.1 2000 ID AAN9300 Combined 904.5 86.9 1060 ID AAN9300 Combined 905.5 86.9 1060 ID AAN9300 ID HEPATIII 905.5 86.9 1060 ID AAN9300 ID HEPATIII 905.5 86.9 1060 ID AAN9300 ID HEPATIII 905.5 8	Score Match Length DB ID  1012 100.0 594 21 AAA7333 98.4 594 21 AAA7333 98.4 594 21 AAA7333 98.4 594 21 AAA7333 98.5 574 21 AAA7333 98.6 2 594 21 AAA7333 98.9 6.7 594 21 AAA7333 99.9 6.8 9 12 10.0 10.0 10.0 10.0 10.0 10.0 10.0 1	Score Match Length DB ID Descripting Match Length DB ID 1015 98 4 594 21 AAA7333 Hepatini 1035 98 4 594 21 AAA7333 Hepatini 998 96.7 594 21 AAA7333 Hepatini 998 96.7 594 21 AAA7333 Hepatini 998 96.7 594 21 AAA7333 Hepatini 973 96.2 594 21 AAA7333 Hepatini 973 96.3 93.3 594 21 AAA7333 Hepatini 973 92.9 588 21 AAA7333 Hepatini 973 92.9 588 21 AAA7333 Hepatini 971 971 971 971 971 971 971 971 971 971	Score Match Length DB ID  1032 100.0 594 21 AAA73330 1035 98.4 594 21 AAA73331 1035 98.4 594 21 AAA73332 98.9 96.7 584 21 AAA73332 98.9 96.7 594 21 AAA73332 98.9 96.7 594 21 AAA73333 98.9 96.7 594 21 AAA73334 98.9 99.2 12734 21 AAA73334 99.9 90.2 12734 21 AAA73334 99.9 90.2 12734 21 AAA73334 90.1 5 87.4 5360 10 AAN900327 90.1 5 87.4 5360 10 AAN900327 90.1 5 87.4 5360 10 AAN900336 90.1 5 87.4 7310 10 AAN900336 90.1 5 87.4 7310 10 AAN900336 90.1 5 87.4 9185 11 AAA72296 90.1 5 87.4 9185 11 AAA72297 90.1 5 87.4 9400 11 AAA72297 90.1 5 87.4 9400 11 AAA72297 90.1 5 87.4 9400 11 AAA72297 90.1 5 87.4 9401 11 AAA72299 90.1 5 87.4 9401 11 AAA72299 90.1 5 87.4 9401 11 AAA72299 90.1 5 87.4 9401 11 AAA72399 90.1 5 87.9 9401 11	Score Match Length DB ID   Descripting	A SCORE MATCH Length DB ID Description of the particular state of the particul	Score Match Length DB   ID   Descripting	Score Match Length DB   ID	Ouery Query Score Match Length DB ID Descripti 1015 98.4 594 21 AAA7333 Hepatiti 1015 98.4 594 21 AAA7333 Hepatiti 1015 98.4 594 21 AAA7333 Hepatiti 1015 98.7 588 21 AAA7333 Hepatiti 1015 98.7 594 21 AAA7333 Hepatiti 1015 99.5 59.4 594 21 AAA7333 Hepatiti 1015 90.5 90.2 12734 594 21 AAA7333 Hepatiti 1015 90.5 90.2 12734 594 21 AAA7333 Hepatiti 1015 90.5 90.2 12734 595 12 AAA7333 Hepatiti 1015 90.5 90.5 97.4 5360 10 AAN92105 Combined 1015 87.4 5360 10 AAN92105 Combined 1015 87.4 7310 10 AAN92105 Composite 1015 87.4 7310 10 AAN92105 Composite 1015 87.4 7310 10 AAN92105 Composite 1015 87.4 9401 13 AAA75296 Composite 1015 87.4 9401 13 AAA73256 Composite 1015 87.4 9401 13 AAA73259 Composite 1015 87.4 9401 13 AAA7333 Hepatiti 1015 87.4 9401 13 AAA73335 Hepatiti 1015 87.4 9401 13 AAA73335 Hepatiti 1015 87.4 9401 13 AAA73359 Hepatiti 1015 87.1 9185 20 AAA26737 Hepatiti 1015 87.1 9185 20	Coore Match Length DB ID    1032   100.0   594   21   AAA73331   Hepatitit   1032   100.0   594   21   AAA73332   Hepatitit   1032   96.7   588   21   AAA73332   Hepatitit   998   96.7   589   21   AAA73333   Hepatitit   991.0   90.2   274   24   AAA73333   Hepatitit   901.5   89.4   21   AAA73334   Hepatitit   901.5   89.4   21   AAA75295   Combined   901.5   89.4   9185   21   AAA75296   Combined   901.5   89.4   9185   21   AAA75296   Combined   901.5   89.4   9185   21   AAA75296   Combined   901.5   89.4   9401   19   AAA79988   Hepatitit   901.5   89.4   9401   19   AAA79998   Hepatitit   901.5   89.4   9401   19   AAA79998   Hepatitit   901.5   89.4   9401   19   AAA79998   Hepatitit   901.5   89.7   9400   19   AAA79998   Hepatitit   901.5   89.7   9400   940	Query Score Match Length DB ID  1032 100.0 594 21 AAA73330 4 594 21 AAA73331 4 594 21 AAA73332 998 96.7 598 21 AAA73332 998 96.7 598 21 AAA73332 998 96.7 598 21 AAA73332 909.5 90.2 12734 24 ABA95163 901.5 87.4 5300 10 AAB95163 901.5 87.4 5300 10 AAB95103 901.5 87.4 5300 10 AAB95103 901.5 87.4 7310 10 AAB95103 901.5 87.4 7310 10 AAB95103 901.5 87.4 7310 10 AAB95103 901.5 87.4 9185 12 AAA7529 901.5 87.4 9185 12 AAA7529 901.5 87.4 9185 12 AAA7529 901.5 87.4 9401 18 AAA7529 901.5	Coore Query  Score Match Length DB ID  1032 100.0 594 21 AAA73330  98.4 594 21 AAA73331  98.6 2 588 21 AAA73333  98.7 594 21 AAA73333  98.9 50.2 588 21 AAA73333  98.9 50.2 588 21 AAA73333  98.9 50.2 588 21 AAA73333  98.9 50.2 12734 24 AAA73333  90.2 12734 24 AAA73333  90.2 12734 24 AAA73333  90.2 12734 24 AAA73334  90.1 5 87.4 5300 10 AAA920505  90.1 5 87.4 5300 10 AAA92037  90.1 5 87.4 7310 10 AAA92037  90.1 5 87.4 7310 10 AAA92035  90.1 5 87.4 7310 10 AAA92035  90.1 5 87.4 9185 12 AAA75297  90.1 5 87.4 9400 13 AAA75296  90.1 5 87.4 9400 13 AAA75297  90.1 5 87.4 9400 13 AAA75297  90.1 5 87.4 9400 13 AAA75297  90.1 5 87.4 9400 14 AAA75297  90.1 5 87.4 9400 17 AAA75297  90.1 5 87.4 9400 18 AAA75297  90.1 5 87.4 9400 18 AAA75297  90.1 5 87.1 9185 20 AAA75249  90.1 5 87.1 9185 20 AAA73334  90.1 5 87.1 9185 20 AAA753259  90.1 5 87.1 9185 20 AAA73334  90.1 5 87.1 9185 20 AAA73344  90.1 5 87.1 9185 20 AAA73344  90.1 5 87.1 9185 20 AAA73344  90.1 5 87.1 9400 10 9400 10 9400 10 9400 10 940	Coore Query Score Match Length DB ID  1032 100.0 594 21 AAA73330 4 20 AAA73333 4 20 AAA73333 5 6 2 594 21 AAA73333 5 6 2 594 21 AAA73333 6 6 3 93.3 594 21 AAA73333 6 9 9 9 9 9 9 9 9 9 9 9 9 9 9 9 9 9 9 9	Coore Query Score Match Length DB ID  1032 100.0 594 21 AAA73330 4 594 21 AAA73333 4 594 21 AAA73333 5 66.2 594 21 AAA73333 5 96.3 594 21 AAA73333 5 96.3 594 21 AAA73333 6 973 94.3 594 21 AAA73334 6 901.5 87.4 5300 10 AAN950105 6 901.5 87.4 5300 10 AAN95106 6 901.5 87.4 7310 10 AAN95106 6 901.5 87.4 7310 10 AAN95106 6 901.5 87.4 9185 11 AAA75296 6 901.5 87.4 9185 12 AAA75296 7 901.5 87.4 9401 18 AA75297 7 901.5 87.4 9401 18 AA75297 7 901.5 87.4 9401 19 AAN99989 7 901.5 87.4 9401 19 AAA799981 7 899 87.1 2058 24 AAA73329 7 899 87.1 2058 24 AAA73259 7 899 87.1 2058 24 AAA73259 7 899 87.1 2058 24 AAA73259 7 897 87.1 9185 20 AAX20555 7 899 87.1 9185 20 AAX20555 7 897 87.1 9185 20 AAX20555 7 899 87.1 9185 20 AAX20555 7 899 87.1 9185 20 AAX20555 7 897 87.1 9185 20 AAX20555 7 897 87.1 9185 20 AAX20555 7 897 87.1 9185 20 AAX20555 7 899 87.1 9185 20 AAX20555 7 897 87.1 9185 20 AAX20555 7 899 87.1 9185 20 AAX20555 7 897 87.1 9185 20 AAX20555 7 899 87.1 9185 2	Coore Query    Score Match Length DB ID   Descripti	Coore Query    Score Match Length DB ID   Descripti	Coore Query  Coore Match Length DB ID  1012 100. 594 21 AAA73330 1015 100. 594 21 AAA73331 1015 100. 594 21 AAA73333 1015 99. 4 594 21 AAA73332 1015 99. 4 594 21 AAA73333 1016 99. 4 594 21 AAA73333 1017 90. 594 21 AAA73333 1018 93. 594 21 AAA73333 1019 930.5 90.2 12734 21 AAA73334 1019 930.5 90.2 12734 21 AAA73334 1019 930.5 90.2 12734 24 ABA95615 1019 930.5 89.3 662 25 AAX15706 1010 901.5 87.4 6905 10 AAN92097 1010 901.5 87.4 6905 10 AAN92003 1011 87.4 7310 10 AAN92033 1011 87.4 9185 11 AA075296 1012 87.4 9185 11 AA075296 1013 87.4 9400 13 AA21744 Hepatiti 1015 87.4 9401 13 AAV21744 Hepatiti 1017 87.4 9401 14 AAN39081 1018 87.4 9401 18 AAV39981 1018 87.4 9401 18 AAV39981 1019 87.1 2058 24 AAA7344 1018 899 87.1 2058 24 AAA7344 1018 899 87.1 2058 24 AAA73259 1018 87.1 87.4 9401 18 AAA7344 1018 899 87.1 2058 24 AAA73259 1018 87.1 87.4 9401 18 AAA7344 1018 899 87.1 2058 24 AAA73259 1018 87.1 87.4 9401 18 AAA73234 1018 899 87.1 2058 24 AAA73259 1018 87.1 87.4 9401 18 AAA73234 1018 899 87.1 2058 24 AAA73259 1018 87.1 87.4 9401 18 AAA73234 1018 899 87.1 2058 24 AAA73234 1018 899 87.1 2058 24 AAA73259 1018 87.1 87.4 AAA73259 1018 87.1 B145 20 AAA73259 1018 87.1 B145 20 AAA732359 1018 87.1 B145 20 AAA732359 1018 87.1 B145 20 AAA732359	Cooker Match Length DB ID    Score Match Length DB ID   1032 100.0   594 21   AAA73330   Hepatiti   1032 100.0   594 21   AAA73333   Hepatiti   998 96.2   594 21   AAA73333   Hepatiti   998 96.3   594 21   AAA73333   Hepatiti   998 96.3   594 21   AAA73334   Hepatiti   991 58 97.4   5360   10   AAN920103   Combined   901 58 97.4   5360   10   AAN920103   Combined   901 58 97.4   7310   16   AA0910336   Combined   901 58 97.4   7310   16   AA0910336   Combined   901 58 97.4   9185   11   AA075297   Compiled   901 58 97.4   9185   12   AA075297   Hepatiti   Hepatiti   901 58 97.4   9401   13   AA021744   Hepatiti   Hepatiti   901 58 97.4   9401   13   AA073043   Hepatiti   Hepatiti   901 58 97.4   9401   13   AA073034   Hepatiti   Hepatiti   901 58 97.4   9401   13   AA073034   Hepatiti   Hepatiti   901 58 97.4   9401   13   AA073034   Hepatiti   901 58 97.4   9401   14   AA073034   Hepatiti   901 58 97.4   9401   14   AA073034   Hepatiti   901 58 97.4   9401   19   AA073034   Hepatiti   901 58 97.4   9401   19   AA073034   Hepatiti   900 58 97.1   2058 24   AAN7344   Hepatiti   900 58 97.1   2058 24   AAN7344   Hepatiti   900 58 97.1   2058 25   AAN7344   Mau2000000000000000000000000000000000000	Ouery Score Mutch Length DB ID Descripti 1032 100.0 594 21 AAA73330 Hepatiti 1012 100.0 594 21 AAA73331 Hepatiti 1012 100.0 594 21 AAA73332 Hepatiti 1012 996 96.7 588 21 AAA73332 Hepatiti 1012 99.0 96.2 594 21 AAA73332 Hepatiti 1012 91.3 594 21 AAA73332 Hepatiti 1012 91.3 594 21 AAA73333 Hepatiti 1012 91.5 81.3 594 21 AAA73334 Hepatiti 1012 91.5 81.3 594 21 AAA73334 Hepatiti 1012 91.5 81.3 51.4 51.2 51.2 51.2 51.2 51.2 51.2 51.2 51.2	Ouery Score Query Score Query Score Match Length DB ID	Coore Query  Coore Match Length DB ID  1013 100.0 594 21 AAA73333  1015 98.4 594 21 AAA73333  4 993 96.2 594 21 AAA73332  4 993 96.2 594 21 AAA73332  4 993 96.2 594 21 AAA73332  5 92.9 588 21 AAA73332  8 92.9 588 21 AAA73332  8 92.9 588 21 AAA73333  9 92.9 588 21 AAA73333  9 92.9 588 21 AAA73332  8 91.5 89.3 612 25 AAA73332  9 91.5 89.4 5300 10 AAN92097  9 91.5 87.4 5300 10 AAN92103  9 91.5 87.4 9185 12 AAA75297  9 901.5 87.4 9185 11 AAO05998  9 901.5 87.4 9400 17 AAA72710  4 901.5 87.4 9401 17 AAA72710  5 901.5 87.4 9401 17 AAA72710  5 901.5 87.4 9401 18 AAA75043  6 901.5 87.4 9401 17 AAA75710  6 901.5 87.4 9401 18 AAA75043  6 901.5 87.4 9401 17 AAA75710  6 901.5 87.4 9401 18 AAA75043  6 901.5 87.4 9401 18	Duery Score Mutch Length DB ID Descripti 1032 100.0 594 21 AAA73330 Hepatiti 998 46.7 594 21 AAA73332 Hepatiti 998 96.7 594 21 AAA73332 Hepatiti 998 96.7 594 21 AAA73332 Hepatiti 993 96.2 594 21 AAA73332 Hepatiti 993 96.2 594 21 AAA73332 Hepatiti 973 94.3 594 21 AAA73332 Hepatiti 973 93.3 594 21 AAA73332 Hepatiti 973 97.9 59.9 588 21 AAA73332 Hepatiti 973 97.5 59.9 59.9 588 21 AAA73332 Hepatiti 97.0 59.0 588 21 AAA73332 Hepatiti 97.0 59.0 588 21 AAA73333 Hepatiti 97.0 59.0 588 21 AAA73334 Hepatiti 97.0 59.0 588 21 AAA73334 Hepatiti 97.0 59.0 588 21 AAA73334 Hepatiti 97.0 58.0 58.0 58.0 58.0 58.0 58.0 58.0 58	Duery Score Query Score Match Length DB ID Descripti 1032 100.0 594 21 AAA73330 Hepatiti 1032 100.0 594 21 AAA73330 Hepatiti 1032 100.0 594 21 AAA73332 Hepatiti 1039 96.2 594 21 AAA73332 Hepatiti 1039 95.4 594 21 AAA73333 Hepatiti 1039 97.9 59.4 594 21 AAA73333 Hepatiti 1039 97.9 59.9 59.9 59.4 21 AAA73333 Hepatiti 1039 97.5 89.3 594 21 AAA73334 Hepatiti 1039 97.5 89.3 594 21 AAA73334 Hepatiti 1039 97.5 89.3 612 25 AAX73334 Hepatiti 1039 97.5 89.4 5300 10 AAN92097 Combined 109.15 87.4 5300 10 AAN92103 Combined 109.15 87.4 7310 10 AAN92103 Combined 109.15 87.4 7310 16 AA082106 Combined 109.15 87.4 9185 11 AA005956 Hepatiti 109.15 87.4 9185 11 AA005956 Sense Store 51 87.4 9185 12 AAA75297 Sense Store 51 87.4 9400 13 AA0110566 Hepatiti 109.15 87.4 9401 18 AA799981 HCV POLY POLY POLY POLY POLY POLY POLY POLY	Cooke Match Length DB ID    Score Match Length DB ID   Descripti	Cooker Match Length DB ID    Score Match Length DB ID   Descripti	Coore Query Score Match Length DB ID  1 1032 100.0 594 21 AAA73330 1 1015 99.6 2 10 AAA73330 1 1015 99.6 2 10 AAA73333 1 1015 96.7 588 21 AAA73333 1 1015 96.2 594 21 AAA73335 1 1015 97.3 94.3 594 21 AAA73335 1 1015 90.2 12734 24 AAA73334 1 1015 90.2 12734 24 AAA73338 1 1015 90.2 12734 24 AAB95615 1 1015 87.4 5360 10 AAN95097 1 1015 87.4 5360 10 AAN95106 1 1015 8	Coore Match Length DB ID    Score Match Length DB ID   Descripti	Coore Query  Coore Match Length DB ID  1032 100.0 594 21 AAA73330 1005 98.4 29 AAA73331 1005 98.4 21 AAA73332 1007 594 21 AAA73332 1008 96.7 598 21 AAA73332 1008 96.7 594 21 AAA73332 1008 96.8 21 AAA73332 1008 96.9 594 21 AAA73332 1008 96.9 1008 1008 1008 1008 1008 1008 1008 100	Duery Score Query Score Match Length DB ID Descripti 1032 100.0 594 21 AAA73330 Hepatiti 1032 100.0 594 21 AAA73330 Hepatiti 1039 96.7 588 21 AAA73333 Hepatiti 1039 96.7 594 21 AAA73335 Hepatiti 1039 96.7 594 21 AAA73334 Hepatiti 1039 96.7 596 97.9 596 97.9 597	Coore Match Length DB ID    Score Match Length DB ID   Descripti     1032 100.0 594 21 AAA73330   Hepatiti     1039 96.7 588 21 AAA73331   Hepatiti     998 96.2 594 21 AAA73335   Hepatiti     999 96.2 594 21 AAA73335   Hepatiti     990 96.2 594 21 AAA73335   Hepatiti     990 96.2 594 21 AAA73335   Hepatiti     990 96.2 594 21 AAA73336     900 96.2 594 21 AAA7336     900 96.2 594 21 AAA736     900 96.2 594 21 AA	Ouery Score Query Score Match Length DB ID Descripti 1032 100.0 594 21 AAA73330 Hepatiti 1015 998 96.7 588 21 AAA7333 Hepatiti 998 96.7 584 21 AAA7333 Hepatiti 993 96.2 594 21 AAA7333 Hepatiti 963 973 94.3 594 21 AAA7333 Hepatiti 953 92.9 588 21 AAA7333 Hepatiti 953 92.9 588 21 AAA7333 Hepatiti 959 92.9 588 3 AAA7332 Hepatiti 959 92.9 588 3 AAA7332 Hepatiti 959 92.9 588 3 AAA7332 Hepatiti 959 92.9 589 3 612 5 AAA7332 Hepatiti 959 90.5 87.4 5300 10 AAN92097 Combined 90.5 87.4 5300 10 AAN92036 Combined 590.5 87.4 5300 10 AAN92106 Combined 590.5 87.4 5310 10 AAN92106 Composite Composite 590.5 87.4 7310 10 AAN92336 Composite Composite 500.5 87.4 7310 10 AAN93336 Composite Composite 500.5 87.4 7310 10 AAN93336 Composite 500.5 87.4 7310 10 AAN93338 Composite 500.5 87.4 7310 10 AAN930338 COMPOSITE 500.5 87.5 87.5 87.5 87.5 87.5 87.5 87.5 87	Coore Match Length DB ID  Score Match Length DB ID  1013 100.0 594 21 AAA73330 1015 98.4 594 21 AAA73331 998 96.7 589 21 AAA73332 998 96.7 594 21 AAA73332 998 96.7 594 21 AAA73333 998 96.8 594 21 AAA73339 998 96.9 594 21 AAA73339 998 96.9 594 21 AAA73333 998 96.9 594 21 AAA73333 998 96.9 594 21 AAA73333 998 96.9 594 21 AAA73334 998 96.9 594 21 AAA73334 998 96.9 588 21 AAA73338 96 96.9 588 21 AAA733338 96 96.9 588 21 AAA73338 96 96 96 96 96 96 96 96 96 96 96 96 96 9	Coore Match Length DB ID    Score Match Length DB ID   Descripti	Score Match Length DB ID    Score Match Length DB ID   Description	Exercise Match Length DB ID    Score Match Length DB ID   Description	Exercised to the control of the cont	Core Match Length DB ID  Score Match Length DB ID  1 1032 100.0 594 21 AAA73330 Hepatitis C v 2 998 96.2 594 21 AAA73335 Hepatitis C v 4 993 96.2 594 21 AAA73335 Hepatitis C v 5 965 95.4 594 21 AAA73332 Hepatitis C v 6 973 94.3 594 21 AAA73332 Hepatitis C v 7 963 93.3 594 21 AAA73332 Hepatitis C v 8 973 94.2 AAA73332 Hepatitis C v 9 973 94.2 AAA73332 Hepatitis C v 9 973 94.2 AAA73332 Hepatitis C v 9 973 973 973 973 973 973 973 973 973 97	Score Match Length DB ID  1 1032 100.0 594 21 AAA73330 Hepatitis C V PS 954 21 AAA73331 Hepatitis C V PS 954 21 AAA73331 Hepatitis C V PS 954 594 21 AAA73335 Hepatitis C V PS 954 594 21 AAA73335 Hepatitis C V PS 954 594 21 AAA73335 Hepatitis C V PS 953 94.3 594 21 AAA73333 Hepatitis C V PS 953 93.3 594 21 AAA73334 Hepatitis C V PS 953 93.3 594 21 AAA73334 Hepatitis C V PS 954 975 965 973 973 973 973 973 973 973 973 973 973	Coore Match Length DB ID    Score Match Length DB ID   Description     1032 100.0 594 21 AAA73330   Hepatitis C     1032 100.0 594 21 AAA73331   Hepatitis C     4 993 96.7 588 21 AAA733329   Hepatitis C     993 96.2 594 21 AAA73332   Hepatitis C     993 96.3 594 21 AAA73333   Hepatitis C     993 96.4 594 21 AAA73333   Hepatitis C     994 96.7 588 21 AAA73333   Hepatitis C     995 96.4 594 594 595     995 96.4 595 9	t Query Score Match Length DB ID Description 1032 100.0 594 21 AAA73330 Hepatitis C 1015 98.4 594 21 AAA73329 Hepatitis C 4 993 96.2 594 21 AAA73329 Hepatitis C Hepatitis C 5 985 95.4 594 21 AAA73332 Hepatitis C 5 985 95.4 594 594 594 594 594 594 594 594 594 59	Score Match Length DB ID  1 1032 100.0 594 21 AAA73330 Hepatitis C 2 1015 98.4 594 21 AAA73331 Hepatitis C 3 998 96.7 588 21 AAA73335 Hepatitis C 4 993 96.2 594 21 AAA73335 Hepatitis C 5 685 954 21 AAA73335 Hepatitis C	t Query  Score Match Length DB ID  1 1032 100.0 594 21 AAA73330 Hepatitis C 2 1015 98.4 594 21 AAA73331 Hepatitis C 3 998 96.7 588 21 AAA73329 Hepatitis C 4 993 96.2 594 21 AAA7335	t Query Score Match Length DB ID 1 1032 100.0 594 21 AAA73330 Hepatitis C 1015 98.4 594 21 AAA73331 Hepatitis C 1015 98.7 588 21 AAA73339 Hepatitis C Hepatitis C 1015 98.7 588 21 AAA73339	Score Match Length DB ID  1 1032 100.0 594 21 AAA73330 Hepatitis C 2 1015 98 4 594 21 AAA73331 Hepatitis C	t Query Score Match Length DB ID 1032 100 0 594 21 abs73330	t Query Coore Match Length DB ID Descripti		CHAMPALES

us-09-965-594-14.rng

301 CCGGCTCCGCAGGGTTCCCGTTCCCTGACCCCGTGCACCTGCGGTTCCTCCGACCTGTAC 360

```
The present sequence is the coding sequence for a mutated version of a fusion protein created using the Hepatitis C virus (HCV) NS3 and NS4A protease enzymes. These proteins are both essential for the replication of the virus, acting to cleave its replicative proteins from the polyprotein produced from the HCV genome. Inhibitors of the two proteins should be effective as antiviral treatments of HCV infection. This is useful as HCV can lead to chronic liver disease such as cirrhosis, liver failure and NS3-NS4A fusion proteins which can be used to identify inhibitors of this type, as well as enabling structural studies of the protease and protease:inhibitor complexes. The protein produced from this sequence contains the alpha-helix0-1 variant.
                                                                                                                                                                                                                                                                                                                                  Modified hepatitis C virus (HCV) NS3 protease comprising at least 1 substitution of a hydrophobic alpha-helix 0 amino acid to a hydrophilic amino acid, useful for screening inhibitors that may treat hepatitis C
                                  /product= "NS4A-NS3 fusion protein #3"
                                                                                                                                                                                                                                               Goldfarb V;
                                                                                                                                                                                                                                               Zhang Y,
                                                                                                                                                                                                                                                                                                                                                                                                                        Claim 26; Fig 13; 66pp; English.
                                                                                                                                                                                                              (BRIM ) BRISTOL-MYERS SQUIBB CO
                                                                                                                                                                                                                                                 Weinheimer S,
                                                                                                                                                                            99US-0115271
                                                                                                                                           2000WO-US00345
                  /*tag=
                                                                                                                                                                                                                                                                                WPI; 2000-465976/40.
P-PSDB; AAB15221.
                                                                     WO200040707-A1
                                                                                                                                                                                                                                                 Wittekind M,
                                                                                                                                           36-JAN-2000;
                                                                                                                                                                            08-JAN-1999;
                                                                                                        13-JUL-2000
```

Sequence 594 BP; 103 A; 186 C; 156 G; 149 T; 0 other;

ProAlaProGlnGlySerArgSerLeuThrProCysThrCysGlySerSerAspLeuTyr 120 TyralaGlnGlnThrArgGlyGluGluGlyCysGlnGluThrSerGlnThrGlyArgAsp LysAsnGlnValGluGlyGluValGlnIleValSerThrAlaAlaGlnThrPheLeuAla ThrCysIleAsnGlyValCysTrpThrValTyrHisGlyAlaGlyThrArgThrIleAla SerProLysGlyProVallleGlnMetTyrThrAsnValAspLysAspLeuValGlyTrp 594 1197 0 0 0 Length:
Matches:
Conservative:
Mismatches:
Indels: US-09-965-594-14 (1-197) x AAA73330 (1-594) 2.42e-85 1032.00 100.00% 100.00% Similarity: Percent Similarity: Alignment Scores: 21 41 121 61 181 241 101 Query Match: Best Local à a ò a ő g õ

40

LeuValThrArgHisAlaAspValIleProValArgArgArgGlyAspSerArgGlySer 140 CIGGITACCCGICACGCIGACGITAICCCGGIICCICGIGGIGGIGACICCCGIGGIICC 420 ProAlaGlyHisAlaValGlyIlePheArgAlaAlaValCysThrArgGlyValAlaLys 180 540 Modified hepatitis C virus (HCV) NS3 protease comprising at least 1 substitution of a hydrophobic alpha-helix 0 amino acid to a hydrophilic amino acid, useful for screening inhibitors that may treat hepatitis C LeuLeuSerProArgProlleSerTyrLeuLysGlySerSerGlyGlyProLeuLeuCys 481 CCGGCTGGTCACGCTGTTGGTATCTTCCGTGCTGTTTGCACCCGTGGTGTTGCTAAA protease; viral replication; chronic liver disease; liver cancer; mutant; mutein; ds. AlavalAspPheIleProValGluSerLeuGluThrThrMetArgSerPro 197 Hepatitis C virus NS4A-NS3 fusion protease coding sequence #4 /\*tag= a /product= "NS4A-NS3 fusion protein #4" Weinheimer S, Zhang Y, Goldfarb V; Location/Qualifiers (BRIM ) BRISTOL-MYERS SQUIBB CO ВÞ 05-JAN-2000; 2000WO-US00345 990S-0115271 AAA73331 standard; DNA; 594 1..594 /\*tag= WPI; 2000-465976/40. Hepatitis C virus. Synthetic. P-PSDB; AAB15222 Hepatitis; NS3 liver failure; WO200040707-A1 Wittekind M, 08-JAN-1999; 19-DEC-2000 13-JUL-2000 AAA73331; 141 161 181 121 361 421 RESULT 2 g ò qq ð g ò 셤

The present sequence is the coding sequence for a mutated version of a fusion protein created using the Hepatitis C virus (HCV) NS3 and NS4A protease enzymes. These proteins are both essential for the replication of the virus, acting to cleave its replicative proteins from the polyprotein produced from the HCV genome. Inhibitors of the two proteins should be effective as antiviral treatments of HCV infection. This is useful as HCV can lead to chronic liver disease such as cirrhosis, liver failure and NS3-NS4A fusion proteins which can be used to identify inhibitors of this type, as well as enabling structural studies of the protease and protease:inhibitor complexes. The protein produced from this sequence contains the alpha-helix0-1 variant. Claim 26; Fig 14; 66pp; English.

180

80

Goldfarb V;

m

```
The present sequence is the coding sequence for a mutated version of a fusion protein created using the Hepatitis C virus (HCV) NS3 and NS4A protease enzymes. These proteins are both essential for the replication of the virus, acting to cleave its replicative proteins from the polyprotein produced from the HCV genome. Inhibitors of the two proteins should be effective as antiviral treatments of HCV infection. This is useful as HCV can lead to chronic liver disease such as cirrhosis, liver failure and Niver cancer. The present invention concerns a number of NS3 mutants and NS3-NS4A fusion proteins which can be used to identify inhibitors of this type, as well as enabling structural studies of the protease and protease:inhibitor complexes. The protein produced from this sequence contains the alpha-helix0-1 variant.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SerProLysGlyProValileGlnMetTyrThrAsnValAspLysAspLeuValGlyTrp 100
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             121 LeuValThrArgHisAlaAspValIleProValArgArgArgGlyAspSerArgGlySer 140
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Modified hepatitis C virus (HCV) NS3 protease comprising at least 1 substitution of a hydrophobic alpha-helix 0 amino acid to a hydrophilic amino acid, useful for screening inhibitors that may treat hepatitis C
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1 MetLysLysLysGlySerValValIleValGlyArgIleAsnLeuSerGlyAspThrAla
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    cceercceaecerrccerrccracccerccaccraccecarrcracaccracrac
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1 ATGAAAAAAAAGGATCCGTTGTTATCGTCGGCCGTATAGTACTGAACGGT-----GCT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                TyralaglnGlnThrargGlyGluGluGlyCysGlnGluThrSerGlnThrGlyArgAsp
                                                                                                                                                                                                                                                                                                                                                                                                                                  BP; 103 A; 180 C; 156 G; 149 T; 0
                                                                                                                               Zhang Y,
                                                                                                                                                                                                                                                    Claim 26; Fig 12; 66pp; English.
                                                                                                        ဗ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         US-09-965-594-14 (1-197) x AAA73329
                                                                                                        (BRIM ) BRISTOL-MYERS SQUIBB
                                                                                                                                                                                                                                                                                                                                                                                                                                                                   3.03e-82
998.00
98.48%
97.97%
                                                                                990S-0115271
                                                                                                                                Weinheimer S,
                                                       06-JAN-2000; 2000WO-US00345
                                                                                                                                                        WPI; 2000-465976/40.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Percent Similarity:
Best Local Similarity:
                                                                                                                                                                    P-PSDB; AAB15220
          WO200040707-A1
                                                                                                                                                                                                                                                                                                                                                                                                                                    Sequence 588
                                                                                                                                Wittekind M,
                                                                                08-JAN-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                          Alignment Scores
                                 13-JUL-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              41
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                21
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      115
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   175
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             81
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     235
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             101
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Query Match:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            à
ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                               140
                                                                                                                                                                                                                                                                             180
                                                                                                                                                                                                                                                                                                                          240
                                                                                                                                                                                                                                                                                                                                                  100
                                                                                                                                                                                                                                                                                                                                                                                                 120
                                                                                                                                                                                                                                                                                                                                                                                                            LeuLeuSerProArgProlleSerTyrLeuLysGlySerSerGlyGlyProLeuLeuCys 160
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      480
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ProAlaGlyHisAlaValGlyIlePheArgAlaAlaValCysThrArgGlyValAlaIyS 180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CCGCTGGTCACGCTGTTGGTATCTTCCGTGCTGCTGTTTGCACCCGTGGTGTTGCTAAA 540
                                                                                                                                                                                                                              120
                                                                                                                                                                                                                                                                                                                                                                         09
                                                                                                                                                                                                                                                                                                   80
                                                                                                                                                                                                      0 4
                                                                                                                                                        20
                                                                                                                                                                              09
                                                                                                                                                                                                                    SerProLysGlyProValIleGlnMetTyThrAsnValAspLysAspLeuValGlyTrp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           LysasnGlnValGluGlyGluValGlnIleValSerThrAlaAlaGlnThrPheLeuAla
                                                                                                                                                                                                                                                                                                  ThrCyslleAsnGlyValCysTrpThrValTyrHisGlyAlaGlyThrArgThrIleAla
                                                                                                                                                                                                                                                                                                                ProAlaProGlnGlySerArgSerLeuThrProCysThrCysGlySerSerAspLeuTyr
                                                                                                                                                                                                                                                                                                                                                                                                                                                LeuValThrArgHisAlaAspValIleProValArgArgArgGlyAspSerArgGlySer
                                                                                                                                                        MetLysLysLysGlySerValValIleValGlyArgIleAsnLeuSerGlyAspThrAla
                                                                                                                                                                    1 ATGAAAAAAAAAGGATCCGTTGTTATCGTCGGCCGTATCAACCTGTCCGGTGACACCGT
                                                                                                                                                                                                        TyralaGlnGlnThrArqGlyGluGluGluGlyCysGlnGluThrSerGlnThrGlyArgAsp
                                                                                                                                                                                                                                                                   Hepatitis; NS3 protease; viral replication; chronic liver disease;
liver failure; liver cancer; mutant; muteln; ds.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             AlaValAspPheIleProValGluSerLeuGluThrThrMetArgSerPro 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Hepatitis C virus NS4A-NS3 fusion protease coding sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              /product= "NS4A-NS3 fusion protein #2"
           other;
                                             594
1194
0
0
0
0
                                                        Matches:
Conservative:
Mismatches:
Indels:
          0
          Ë
          G; 147
                                             Length:
                                                                                                                                US-09-965-594-14 (1-197) x AAA73331 (1-594)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Location/Qualifiers
          155
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    вР
           ت
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    AAA73329 standard; DNA; 588
          A; 187
                                             8.61e-84
1015.00
98.48%
98.48%
98.35%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 19-DEC-2000 (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ø
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   /*tag=
          BP; 105
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Hepatitis C virus.
                                                                                 Similarity:
                                                                      Percent Similarity:
          594
                                   Alignment Scores
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Synthetic
            Sequence
                                                                                                                                                                                                        21
                                                                                                                                                                                                                              61
                                                                                                                                                                                                                                                     41
                                                                                                                                                                                                                                                                             121
                                                                                                                                                                                                                                                                                                   61
                                                                                                                                                                                                                                                                                                                          181
                                                                                                                                                                                                                                                                                                                                                  81
                                                                                                                                                                                                                                                                                                                                                                          241
                                                                                                                                                                                                                                                                                                                                                                                                 101
                                                                                                                                                                                                                                                                                                                                                                                                                        301
                                                                                                                                                                                                                                                                                                                                                                                                                                                121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                       361
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               141
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       421
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               161
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     481
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            AAA73329
                                                                                Best Local Si
Query Match:
DB:
                                               Pred. No.:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                       đ
                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                  ద
                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      δ
                                                                                                                                                        à
                                                                                                                                                                             a
                                                                                                                                                                                                      ò
                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                ð
                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                  Q
                                                                                                                                                                                                                                                                                                                     q
                                                                                                                                                                                                                                                                                                                                             οχ
```

354

174

9

40

20 54

588 1193 1

Conservative: Mismatches: Indels:

(1-588)

Length: Matches:

other

234

AAA73335

q õ a

õ

```
Conservative:
Mismatches:
                                     Matches:
                Length:
                                                                                          Indels:
                                                                                                                                               US-09-965-594-14 (1-197) x AAA73335 (1-594)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Location/Qualifiers
1..594
/*tag= a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          BP.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          AAA73332 standard; DNA; 594
                    8.79e-82
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (first entry)
                                 993.00
97.46%
97.46%
96.22%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Hepatitis C virus.
Synthetic.
                                                    Percent Similarity:
Best Local Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               19-DEC-2000
 Alignment Scores
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     361
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         141
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            161
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                481
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              AAA73332;
                                                                                          Query Match:
                Pred. No.:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        RESULT 5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       AAA73332
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Ω
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                    ô
                                                                                                                                                                                                                                                                                      a
                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                  a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ö
                                                                                                                                                                                 ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              The present sequence is the coding sequence for a mutated version of a fusion protein created using the Hepatitis C virus (HCV) NS3 and NS4A protease enzymes. These proteins are both essential for the replication of the virus, acting to cleave its replicative proteins from the polyprotein produced from the HCV genome. Inhibitors of the two proteins should be effective as antiviral treatments of HCV infection. This is useful as HCV can lead to chronic liver disease such as cirrhosis, liver failure and liver cancer. The present invention concerns a number of NS3 mutants and NS3-NS4A fusion proteins which can be used to identify protease and protease:inhibitors complexes. The protein produced from this sequence contains the alpha-helixO wild-type sequence.
                                                                                                                                             ProAlaGlyHisAlaValGlyIlePheArgAlaAlaValCysThrArgGlyValAlaLys 180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Modified hepatitis C virus (HCV) NS3 protease comprising at least 1 substitution of a hydrophobic alpha-helix 0 amino acid to a hydrophilic amino acid, useful for screening inhibitors that may treat hepatitis C
                                                    CTGGTTACCCGTCACGCTGACGTTATCCCGGTTCGTCGTCGTGGTGACTCCCGTGGTTCC
Hepatitis; NS3 protease; viral replication; chronic liver disease;
liver failure; liver cancer; mutant; mutein; ds.
                                                                                                                                                                                                Hepatitis C virus NS4A-NS3 fusion protease coding sequence #8.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 protein #8"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Sequence 594 BP; 98 A; 189 C; 153 G; 154 T; 0 other;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Goldfarb
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 "NS4A-NS3 fusion
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Zhang Y,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Location/Qualifiers
1..594
/*tag= a
/product= "NS4A-NS3"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Disclosure; Fig 18; 66pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ပ္ပ
                                                                                                                                                                                                                                                                                                           ВР
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  (BRIM ) BRISTOL-MYERS SQUIBB
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Wittekind M, Weinheimer S,
                                                                                                                                                                                                                                                                                                           AAA73335 standard; DNA; 594
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         2000WO-US00345
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              99US-0115271
                                                                                                                                                                                                                                                                                                                                                                                 (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  WPI; 2000-465976/40.
P-PSDB; AAB15226.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     WO200040707-A1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         06-JAN-2000;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              08-JAN-1999;
                                                                                                                                                                                                                                                                                                                                                                               19-DEC-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         13-JUL-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Synthetic
                                                                                                                                                                                                                                                                                                                                             AAA73335;
                                                                                        415
                                                      141
```

SXXXBXBXBXSSXEFFFFXXBXFXFXBXFXBXBXSSXXBXBXBXSSXXBXBXBXSSXSXBXBXSXBX

```
140
                                                                                                                                                                                                                                                                                                                            420
                                                                                                                                                                                                                                                                                                                                               160
                                                                                                                                                                                                                                                                                                                                                                480
                                                                                                                                                                                                                                                                                                                                                                                                      540
                                                                                                                                                                           180
                                                                                                                                                                                                                                  100
                                                                                                                                                                                                                                                    300
                                                                                                                                                                                                                                                                                        360
                                                                                                                                                                                                                                                                                                                                                                                   ProAlaGlyHisAlaValGlyIlePheArgAlaAlaValCysThrArgGlyValAlaLys 180
                                                                                                                                                                                              80
                                                                                                   9
                                                                                                                     40
                                                                                                                                                          9
                                                                                                                                                                                                                                                                                                                                                LeuLeuSerProArgProlleSerTyrLeuLysGlySerSerGlyGlyProLeuLeuCys
                                                                                 CIGCIGICCCGGGGICCGAICTCCTACCTGAAAGGTICCTCCGGGGGGCGGTGCTGGTGG
                                                                                                                                                                                                                                                                                                                                                                                                       CCGGCTGGTCACGCTGTTGGTATCTTCCGTGCTGCTGTTTGCACCCGTGGTGTTGCTAAA
                                                                                                                     21 TyralaGlnGlnThrArgGlyGluGluGlyCysGlnGluThrSerGlnThrGlyArgAsp
                                                                                                                                                           41 LysAsnGlnValGluGlyGluValGlnIleValSerThrAlaAlaGlnThrPheLeuAla
                                                                                                                                                                   LeuValThrArgHisAlaAspValIleProValArgArgArgGlyAspSerArgGlySer
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Hepatitis; NS3 protease; viral replication; chronic liver disease;
liver failure; liver cancer; mutant; mutein; ds.
                                                                                                                                                                                                                                                                                                                                                                                                                       Hepatitis C virus NS4A-NS3 fusion protease coding sequence #5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  /*tag= a
/product= "NS4A-NS3 fusion protein #5"
Gaps:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             WO200040707-A1
```

S

```
Synthetic
                              421
                                                                                          481
                                                             161
                                                                                                                        à
                             g
                                                             ô
                                                                                        a
                                                                                                                                                                                                                                                                                                                   fusion protein created using the Hepatitis C virus (HCV) NS3 and NS4A protease enzymes. These proteins are both essential for the replication of the virus, acting to cleave its replicative proteins from the polyprotein produced from the HCV genome. Inhibitors of the two proteins should be effective as antiviral treatments of HCV infection. This is should be effective as antiviral treatments of HCV infection. This is useful as HCV can lead to chronic liver disease such as cirrhosis, liver failure and liver cancer. The present invention concerns a number of NS3 mutants and NS3-NS4A fusion proteins which can be used to identify inhibitors of this type, as well as enabling structural studies of the protease and protease:inhibitor complexes. The protein produced from this sequence contains the alpha-helix0-1 variant.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     100
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               40
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       80
                                                                                                                                                                                                 Modified hepatitis C virus (HCV) NS3 protease comprising at least 1 substitution of a hydrophobic alpha-helix 0 amino acid to a hydrophilic amino acid, useful for screening inhibitors that may treat hepatitis C
                                                                                                                                                                                                                                                                                                             present sequence is the coding sequence for a mutated version of a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          21 TyralaGlnGlnThrArgGlyGluGluGlyCysGlnGluThrSerGlnThrGlyArgAsp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ACCTCCATCAACGGTGTTCTGTGGACCGTTTACCACGGTGCTGGTACCCGTACCATCGCT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SerProLysGlyProValIleGlnMetTyrThrAsnValAspLysAspLeuValGlyTrp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   MetLysLysLysGlySerValValIleValGlyArgIleAsnLeuSerGlyAspThrAla
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          LysAsnGlnValGluGlyGluValGlnIleValSerThrAlaAlaGlnThrPheLeuAla
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ThrCys1leAsnGlyValCysTrpThrValTyrHisGlyAlaGlyThrArgThrIleAla
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                other;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           594
191
0
6
0
                                                                                                                        Goldfarb V;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Length:
Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Sequence 594 BP; 105 A; 189 C; 153 G; 147 T; 0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (1-594)
                                                                                                                        Zhang Y,
                                                                                                                                                                                                                                                                              Claim 26; Fig 15; 66pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      US-09-965-594-14 (1-197) x AAA73332
                                                                                          (BRIM ) BRISTOL-MYERS SQUIBB
                                                                                                                        Weinheimer S,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        4.72e-81
985.00
96.95%
96.95%
                                2000WO-US00345
                                                             99US-0115271
                                                                                                                                                     WPI; 2000-465976/40.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Percent Similarity:
Best Local Similarity:
                                                                                                                                                                       P-PSDB; AAB15223
                                                                                                                        Wittekind M,
                              06-JAN-2000;
                                                             08-JAN-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Alignment Scores:
13-JUL-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    81
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   83
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Query Match:
DB:
ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                엄
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
```

```
inhibitors of this type, as well as enabling structural studies of the protease and protease:inhibitor complexes. The protein produced from this sequence contains the alpha-helix0-7 variant.
                                                                  ProAlaGlyHisAlaValGlyIlePheArgAlaAlaValCysThrArgGlyValAlaLys 180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            fusion protein created using the Hepatitis C virus (HCV) NS3 and NS4A protease enzymes. These proteins are both essential for the replication of the virus, acting to cleave its replicative proteins from the polyprotein produced from the HCV genome. Inhibitors of the two proteins should be effective as antiviral treatments of HCV infection. This is useful as HCV can lead to chronic liver disease such as cirrhosis, liver failure and liver cancer. The present invention concerns a number of NS3 mutants and NS3-NS4A fusion proteins which can be used to identify
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Modified hepatitis C virus (HCV) NS3 protease comprising at least 1 substitution of a hydrophobic alpha-helix 0 amino acid to a hydrophilic amino acid, useful for screening inhibitors that may treat hepatitis C
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                The present sequence is the coding sequence for a mutated version of a
                 141 LeuLeuSerProArgProlleSerTyrLeuLysGlySerSerGlyGlyProLeuLeuCys
                                                                                                                                                                                                                                                                                                                                                                            Hepatitis; NS3 protease; viral replication; chronic liver disease;
liver failure, liver cancer; mutant; mutein; ds.
                                                                                                                                  Hepatitis C'virus NS4A-NS3 fusion protease coding sequence #6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1..594
/*tag= a
/product= "NS4A-NS3 fusion protein #6"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Sequence 594 BP; 104 A; 191 C; 152 G; 147 T; 0 other;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 594
188
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Goldfarb
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Wittekind M, Weinheimer S, Zhang Y,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Claim 26; Fig 16; 66pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          (BRIM ) BRISTOL-MYERS SQUIBB
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 5.88e-80
973.00
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        990S-0115271
                                                                                                                                                                                                                                        AAA73333 standard; DNA; 594
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        06-JAN-2000; 2000WO-US00345
                                                                                                                                                                                                                                                                                                           (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            WPI; 2000-465976/40.
                                                                                                                                                                                                                                                                                                                                                                                                                              Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            P-PSDB; AAB15224
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    WO200040707-A1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         08-JAN-1999;
                                                                                                                                                                                                                                                                                                          19-DEC-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Alignment Scores:
Pred. No.:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      13-JUL-2000
                                                                                                                                                                                                                                                                          ProAlaProGlnGlySerArgSerLeuThrProCysThrCysGlySerSerAspLeuTyr 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    301 CAGGCTCCGCAGGGTTCCCGTTCCCTGACCCCGTGCACCTGCGGTTCCTCCGACCTGTAC
```

241

101

ò đ

```
180
                                                                                                                                                                                                                                                                                                                                                    540
                                                                                                                                                                                240
                                                                                                                                                                                               SerProLysGlyProValileGlnMetTyrThrAsnValAspLysAspLeuValGlyTrp 100
                                                                                                                                                                                                                                                                   140
                                                                                                                                                                                                                                                                                  420
                                                                                                      9
                                                                                                                                                               80
                                                                             9
                                                                                                                                                                                                                                                                                                                                   CCGGCTGGTCACGCTGTTGGTATCTTCCGTGCTGTTTCCACCCGTGGTGTTGCTAAA
                                                                                                                                                                                                                                                                  LeuValThrArgHisAlaAspValIleProValArgArgArgGlyAspSerArgGlySer
                                                          LysasnGlnValGluGlyGluValGlnIleValSerThrAlaAlaGlnThrPheLeuAla
                                                                                                                                        ThrCysIleAsnGlyValCysTrpThrValTyrHisGlyAlaGlyThrArgThrIleAla
                                                                                                                                                                         CTGGTTACCCGTCACGCTGACGTTATCCCGGTTCGTCGTCGTCGTGGTGACTCCCGTGGTTAC
                                                                                             TyralaglnGlnThrargGlyGluGluGlyCysGlnGluThrSerGlnThrGlyArgAsp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Hepatítis; NS3 protease; viral replication; chronic liver disease;
liver failure; liver cancer; mutant; mutein; ds.
                                                                                                                                                                                                                                                                                                                                                                     AlaValAspPheIleProValGluSerLeuGluThrThrMetArgSerPro 197
                                                                                                                                                                                                                                                                                                                                                                             Hepatitis C virus NS4A-NS3 fusion protease coding sequence #7.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     /product= "NS4A-NS3 fusion protein #7"
 2000
 Conservative:
        Mismatches:
Indels:
                          Gaps:
                                           (1-594)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Location/Qualifiers
1..594
                                           US-09-965-594-14 (1-197) x AAA73333
                                                                                                                                                                                                                                                                                                                                                                                                                       ВР
                                                                                                                                                                                                                                                                                                                                                                                                                       AAA73334 standard; DNA; 594
                                                                                                                                                                                                                                                                                                                                                                                                                                                        19-DEC-2000 (first entry)
96.45%
95.43%
94.28%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              /*tag=
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  virus
         Best Local Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      WO200040707-A1
Percent Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Hepatitis C v
Synthetic.
                                                                            ٦
                                                                                                                                                                                                                 241
                                                                                                                                                                                                                                                                                                                                                  481
                                                                                                                                                                                                                                                                                                                                                                                     541
                                                                                                              61
                                                                                                                              4.1
                                                                                                                                                               61
                                                                                                                                                                                                                                                                   121
                                                                                                                                                                                                                                                                                   361
                                                                                                                                                                                                                                                                                                   141
                                                                                                                                                                                                                                                                                                                   421
                                                                                                                                                                                                                                                                                                                                    191
                                                                                                                                                                                                                                                                                                                                                                                                                                      AAA73334;
                                                                                                                                                                                                81
                                                                                                                                                                                                                                                                                                                                                                    181
                   Query Match:
                                                                                                                                                                                                                                                                                                                                                                                                       RESULT 7
                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                            g
                                                                                               ö
                                                                                                             셤
                                                                                                                              à
                                                                                                                                              a
                                                                                                                                                               ô
                                                                                                                                                                              q
                                                                                                                                                                                                ò
                                                                                                                                                                                                                g
                                                                                                                                                                                                                                  ö
                                                                                                                                                                                                                                                 a
                                                                                                                                                                                                                                                                   ç
                                                                                                                                                                                                                                                                                  셤
                                                                                                                                                                                                                                                                                                  ö
                                                                                                                                                                                                                                                                                                                  a
                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                    셤
                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                     셤
```

06-JAN-2000; 2000WO-US00345

```
240
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       101 ProAlaProGlnGlySerArgSerLeuThrProCysThrCysGlySerSerAspLeuTyr 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                SerProLysGlyProVallleGlnMetTyrThrAsnValAspLysAspLeuValGlyTrp 100
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                301 CAGGCTCCGCAGGGTTCCCTGACCCCGTGCACCTGCGGTTCCTCCGACCTGTAC 360
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   LeuValThrArgHisAlaAspValIleProValArgArgArgGlyAspSerArgGlySer 140
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 9
                                                                                                                                                                                                                                      The present sequence is the coding sequence for a mutated version of a fusion protein created using the Hepatitis C virus (HCV) NS3 and NS4A protease enzymes. These proteins are both essential for the replication of the virus, acting to cleave its replicative proteins from the polyprotein produced from the HCV genome. Inhibitors of the two proteins should be effective as antiviral treatments of HCV infection. This is useful as HCV can lead to chronic liver disease such as cirrhosis, liver failure and liver cancer. The present invention concerns a number of NS3 mutants and NS3-NS4A fusion proteins which can be used to identify inhibitors of this type, as well as enabling structural studies of the protease and protease:inhibitor complexes. The protein produced from this sequence contains the alpha-helix0-7 variant.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    20
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ThrCyslleAsnGlyValCysTrpThrValTyrHisGlyAlaGlyThrArgThrIleAla 80
                                                                                                                                         Modified hepatitis C virus (HCV) NS3 protease comprising at least 1 substitution of a hydrophobic alpha-helix 0 amino acid to a hydrophillc amino acid, useful for screening inhibitors that may treat hepatitis C
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ACCTCCATCAACGGTGTTCTGTGGACCGTTTACCACGGTGCTGGTACCCGTACCATCGCT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  LysasnGlnValGluGlyGluValGlnIleValSerThrAlaAlaGlnThrPheLeuAla
                                                                                                                                                                                                                                                                                                                                                                                                                               other;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                       594
187
2
0
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Conservative:
Mismatches:
Indels:
                                                                      Goldfarb
                                                                                                                                                                                                                                                                                                                                                                                                                             Sequence 594 BP; 105 A; 192 C; 151 G; 146 T; 0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Length:
Matches:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Gaps:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       US-09-965-594-14 (1-197) x AAA73334 (1-594)
                                                                      Zhang Y,
                                         8
                                                                                                                                                                                                                Claim 26; Fig 17; 66pp; English.
                                         (BRIM ) BRISTOL-MYERS SQUIBB
                                                                      ò
             99US-0115271
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  963.00
95.94%
94.92%
93.31%
                                                                      Weinheimer
                                                                                                 WPI; 2000-465976/40
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Similarity:
                                                                                                              P-PSDB; AAB15225
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Percent Similarity:
             08-JAN-1999;
                                                                    Wittekind M,
                                                                                                                                                                                                                                                                                                                                                                                                                                                           Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 41
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 81
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                361
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           141
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      421
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              241
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Query Match:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Best Local
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    q
ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 οŽ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      셤
```

31 15:59:14 2003

Sun Aug

```
g
                                                                                                                                                                                                                                                                                                                                                                                                                셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              οy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      셤
                                                                            ò
                                                                                                   qq
                                                                                                                               ò
                                                                                                                                                       g
                                                                                                                                                                               ò
                                                                                                                                                                                                        g
                                                                                                                                                                                                                                  õ
                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         The present sequence is the coding sequence for a fusion protein created using the Hepatitis C virus (HCV) NS3 and NS4A protease enzymes. These proteins are both essential for the replication of the virus, acting to cleave its replicative proteins from the polyprotein produced from the HCV genome. Inhibitors of the two proteins should be effective as antiviral treatments of HCV infection. This is useful as HCV can lead to chronic liver disease such as cirrhosis, liver failure and liver cancer. The present invention concerns a number of NS3 mutants and NS3-NS4A fusion proteins which can be used to identify inhibitors of this type, as
                ProAlaClyHisAlaValGlyIlePheArgAlaAlaValCysThrArgGlyValAlaLys 180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Modified hepatitis C virus (HCV) NS3 protease comprising at least 1 substitution of a hydrophobic alpha-helix 0 amino acid to a hydrophilic amino acid, useful for screening inhibitors that may treat hepatitis C
                                                                                                                                                                                                                           protease; viral replication; chronic liver disease; liver cancer; ds.
                                                    AlaValAspPheIleProValGluSerLeuGluThrThrMetArgSerPro 197
                                                              Hepatitis C virus NS4A-NS3 fusion protease coding sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              well as enabling structural studies of the protease and
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          588
188
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   BP; 97 A; 183 C; 153 G; 155 T; 0 other;
                                                                                                                                                                                                                                                                                                                                                 /product= "NS3-NS4A fusion protein"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Goldfarb V;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Matches:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Length:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Wittekind M, Weinheimer S, 2hang Y,
                                                                                                                                                                                                                                                                                                             Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Disclosure; Fig 10; 66pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ဗ
                                                                                                                               Bb
                                                                                                                                                                                                                                                                                                                                                                                                                                                                               (BRIM ) BRISTOL-MYERS SQUIBB
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          protease:inhibitor complexes
                                                                                                                               AAA73328 standard; DNA; 588
                                                                                                                                                                                                                                                                                                                                                                                                                                                      99US-0115271
                                                                                                                                                                                                                                                                                                                                                                                                                              06-JAN-2000; 2000WO-US00345
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1.1e-78
959.00
95.94%
95.43%
92.93%
                                                                                                                                                                               (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  WPI; 2000-465976/40.
                                                                                                                                                                                                                                                                   Hepatitis C virus.
Synthetic.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Percent Similarity:
Best Local Similarity:
Query Match:
DB:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               P-PSDB; AAB15212
                                                                                                                                                                                                                                  Hepatitis; NS3
liver failure;
                                                                                                                                                                                                                                                                                                                                                                           WO200040707-A1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Sequence 588
                                                                                                                                                                                                                                                                                                                                                                                                                                                      08-JAN-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Alignment Scores:
                                                                                                                                                                               19-DEC-2000
                                                                                                                                                                                                                                                                                                                                                                                                    13-JUL-2000
                          481
                                                  181
161
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Pred. No.:
```

```
140
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            234
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     100
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  294
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                414
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             161 ProAlaGlyHisAlaValGlyIlePheArgAlaAlaValCysThrArgGlyValAlaLys 180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            9
                                                                                                                                                                                                                                                                                                                                                                                                                                                    80
                                                                                                                                                                                          40
                                                                                                                             54
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1 MetLysLysLysGlySerValValileValGlyArgileAsnLeuSerGlyAspThrAla
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SerProLysGlyProVallleGlnMetTyrThrAsnValAspLysAspLeuValGlyTrp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1 ATGAAAAAAAAGGTTCCGTTGTTATCGTCGGCCGTATAGTACTGAACGGT-----GCT
                                                                                                                                                                                          21 TyralaGlnGlnThrArqGlyGluGluGlyCysGlnGluThrSerGlnThrGlyArqAsp
                                                                                                                                                                                                                                                                                                                       41 LysAsnGlnValGluGlyGluValGlnIleValSerThrAlaAlaGlnThrPheLeuAla
                                                                                                                                                                                                                                                                                                                                                                                                                                                    61 ThrCyslleAsnGlyValCysTrpThrValTyrHisGlyAlaGlyThrArgThrlleAla
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            181 AlaValAspPhelleProValGluSerLeuGluThrThrMetArgSerPro 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Nucleic acid construct encoding chimeric Hepatitis C Virus (HCV)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Pestivirus; Npro; protease; NS3; screening; ds.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Chimeric - Bovine viral diarrhea virus.
Chimeric - Hepatitis C virus.
US-09-965-594-14 (1-197) x AAA73328 (1-588)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Chimeric BVDV/HCV NS3-wt sequence.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ВР
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ABA95615 standard; DNA; 12734
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      JYN;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  99US-0344456
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               99US-0344456
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Lai VCH, Lau
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (SCHE ) SCHERING CORP.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                WPI; 2002-121103/16.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 US6326137-B1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  25-JUN-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               25-JUN-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     21-MAR-2002
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  04 - DEC - 2001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 81
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ABA95615;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Hong Z,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           RESULT 9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ABA95615
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           THE STATE OF THE S
```

```
protease gene, useful for in vivo screening of compounds which inhibit HCV infection
pestivirus genome where the Npro protease gene is replaced with NS3
```

Example 2; Columns 17-28; 20pp; English.

Anti-viral synthetic prototoxophore associated DNA sequence

(first entry)

28-MAR-2003

chimeric Hepatitis C virus (HCV)-pestivirus genome. The construct comprises a pestivirus genome where a Npro pestivirus protease gene is replaced with a gene encoding a functional HCV NS3 protease. Furthermore, each junction site recognised by the Npro protease is replaced with a junction site recognised by the HCV NS3 protease. The construct is useful for screening compounds that inhibit HCV in vivo by inhibiting HCV protease, where screening may be in cell culture or in an animal model. The present sequence is a chimmeric clone of BVDV (bovine viral diarrhea virus)/HCV NS3-wt, which was used to illustrate the present invention. present invention relates to a nucleic acid construct encoding a

Sequence 12734 BP; 4032 A; 2604 C; 3295 G; 2803 T; 0 other;

82 ProLysGlyProVallleGlnMetTyrThrAsnValAspLysAspLeuValGlyTrpPro 101 ValThrArgHisAlaAspValIleProValArgArgArgGlyAspSerArgGlySerLeu 141 LeuSerProArgProlleSerTyrLeuLysGlySerSerGlyGlyProLeuLeuCysPro 161 CTITCGCCCCGGCCCATTTCCTACCTAAAAGGCTCCTCGGGGGGGTCCGCTGTTGTGCCCC 892 162 AlaGlyHisAlaValGlyIlePheArgAlaAlaValCySThrArgGlyValAlaLySAla 181 GlyServalvalIlevalGlyArgIleAsnLeuSerGlyAsp-----ThrAlaTyr 22 AlaGlnGlnThrArgGlyGluGluGlyCysGlnGluThrSerGlnThrGlyArgAspLys 42 AsnGlnValGluGlyGluValGlnIleValSerThrAlaAlaGlnThrPheLeuAlaThr Conservative: Mismatches: Indels: Matches: Length: (1-12734)Gaps: US-09-965-594-14 (1-197) x ABA95615 1.94e-74 930.50 94.87% 92.82% Percent Similarity: Best Local Similarity: Alignment Scores: 122 773 142 S Query Match: Score: g õ 셤 ö ద õ 셤 g ò õ ö ð ö 셤 ö g 셤

```
This invention relates to a novel synthetic viral prototoxophore domains specific for a viral enzyme. This prototoxophore may be bound and and specific for a viral enzyme. This prototoxophore may be bound and modified by the viral enzyme thus converting it to a toxophore. Also disclosed in the invention is a method for enhancing the anti-viral entry of its susceptible to infection, with a cell, coffected with a virus or is susceptible to infection, with a prototoxophore. The invention further comprises an assay to identify a nati-viral agents, comprising contacting an infected cell with a candidate agent and comparing the ability of the agent to inhibit the growth or infectivity of the virus in the cell. The prototoxophores of the invention may have viruside or antitumour activity. The prototoxophores of the invention may be useful for reducing or inhibiting viral infectivity, by contacting a cell (e.g. lymphocyte, onerve cell, connective tissue cell, muscle cell or hepotocyte) which is infected with a virus or is susceptible to infection with a virus, with a cadapted to long term continuous culture or isolated from a subject. The prototoxophore is also useful for ameliorating the severity of a viral infection in a subject, where the virus is selected from human immunodeficiency virus (Hist), herpes simplex virus (HSV), rhinovirus and hepatitis virus, by administering an effective amount of the prototoxophore to the subject. The prototoxophores of the invention are antiviral prototoxophore associated DNA sequence represents an antiviral prototoxophore associated DNA sequence, this sequence is a networthy and received by administering and prototoxophore is a recombinant NASI/NS4 fusion protein and antiviral prototoxophore is a recombinant prototoxophore is a re
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    has
                                                                                                                                                 Hepatitis C; ds; viral prototoxophore; anti-viral; tumour; virus; infection; antitumour; toxophore; human immunodeficiency virus; HIV infection; herpes simplex virus; HSV; rhinovirus; NS3 protease.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Novel synthetic viral prototoxophore for treating viral infections, toxin molety incorporated into substrate domain specific for viral enzyme, bound and modified by viral enzyme to get converted into toxophore
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      the invention although it is clearly not a protein sequence.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ₹
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Shepard
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Example 1; Page 62; 66pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                           26-APR-2002; 2002WO-US13223
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Neuteboom STC,
                                                                                                                                                                                                                                                                                                                                                                                                                                                        27-APR-2001; 2001US-286893P
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          (NEWB-) NEWBIOTICS INC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            WPI; 2003-167102/16.
                                                                                                                                                                                                                                                                                                       WO200287500-A2
                                                                                                                                                                                                                                                     Unidentified
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Cathers BE,
                                                                                                                                                                                                                                                                                                                                                         07-NOV-2002
```

532

47

592

61

652

81

612 181 3 8 8 1

Length:
Matches:
Conservative:
Mismatches:
Indels:

921.50 94.36% 92.82% 89.29%

Percent Similarity: Best Local Similarity:

Query Match:

(1-612)

US-09-965-594-14 (1-197) x ABX15706

Sequence 612 BP; 120 A; 171 C; 191 G; 130 T; 0 other

.07e-75

Alignment Scores:

182 ValAspPheIleProValGluSerLeuGluThrThrMetArgSer 196 

BP.

ABX15706 standard; DNA; 612

953

8

à

RESULT 10 ABX15706 σ

```
It is a double-stranded nucleotide sequence of the open reading frame (ORF) (tag a) extending through clones 14i, 11b, 7f, 7e, 8h 33c, 40b, 31b, 35, 36, 11, 32, 33b, 25c, 14c, 8f, 31f, 31g and 39c of hepatitis C virus (HCV) cDNA. In creating the composite sequence the following heterogeneities were considered. Clone 33c contains a sequence of 800 base pairs which overlaps the conneas 40b and 37c. In clone 31c, as well as in 5 other overlapping clones, nucleotide #789 is a G. However, in clone 37b the corresponding nucleotide is an A. This chertogenity may have important ramifications for protein folding. Concloses the corresponding residue in clone 7e and in 3 other overlapping clones is an A. Therefore the residue in the 7e and in 3 other overlapping clones is an A. Therefore the residue in the 3 other overlapping clones is an A. Therefore the residue in clone 31f and in 2 other overlapping clones is an A. The 3' terminal sequence of clone 31f is sequence of clone 31f is sequence 1 in clone 31f and in 2 other overlapping clones, rather than as TrGC, as is found in the corresponding residue in clone 31f and 2 other overlapping clones is an A. The 3' terminal sequence of clone 31f is sequence corresponding residue in clone 31f and 2 other overlapping clones is an A. The 3' terminal sequence and a rather than a TrGC clones is an A. The 3' terminal clone 31f and 2 other overlapping clones is corresponding sequences in none 31f and 2 other overlapping clones is an A. The 3' terminal residue in clone 31f and 3 other clones is an A. The 3' terminal response to depicted as TA rather than a Connesting and the corresponding sequences in none 31f and 2 other overlapping clones are shown. AAN02097 could be used as a source of clones is TA. Potential cloning artifacts have been omitted and instead overlapping clones are shown. AAN02097 could be used as a source of clones is and passive immunotherany for HCV infarthorm.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   927 TIGCIGGCCCCATCACGCCTACGCCCAGCAGACAAGGGGCCTCCTAGGGTGCATAATC 986
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           15 ---LeuSerGlyAspThrAlaTyrAlaGlnGlnThrArgGlyGluGluGlyCysGlnGlu 33
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              34 ThrSerGlnThrGlyArgAspLySAsnGlnValGluGlyGluValGln1leValSerThr
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        867 CGCAGGGCCCGGGAGATACTGCTCGGGCCAGCCGATGGAATGGTCTCCAAGGGGTGGAGG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 387 ACCAGCCTAACTGGCCGGGACAAAACCAAGTGGAGGTGAGGTCCAGATTGTGTGTCAACT
                                                                                                                                                                                           Claim 3; Figure 26-1, 26-2, 26-3, 26-4, 26-5, 26-6; 139pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     lmmuno- assay reagents and vaccines and to generate antibodies used
diagnosis and passive immunotherapy for HCV infection/non-A, non-B
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Sequence 5300 BP; 1047 A; 1606 C; 1515 G; 1130 T; 2 other;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       3 LysLysGlySerValVallleValGlyArglleAsn...---
                                                                                                                                        Purified hepatitis C virus - and associated nucleic acids and polypeptide(s)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        (Updated on 25-MAR-2003 to correct PR field.) (Updated on 25-MAR-2003 to correct PI field.)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Length:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     US-09-965-594-14 (1-197) x AAN92097 (1-5300)
                                                        Ö
                                                          Kuo
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              2.93e-72
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             901.50
90.20%
85.78%
87.35%
                                                        Choo QL,
                     (CHIR ) CHIRON CORP
                                                                                         WPI; 1989-159274/22
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Similarity:
                                                                                                           P-PSDB; AAP92041
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Percent Similarity:
Best Local Similari
                                                        Houghton M,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           hepatitis.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Query Match:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          q
       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                       ProLysGlyProValileGlnMetTyrThrAsnValAspLysAspLeuValGlyTrpPro 101
                                                                                       LeuSerProArgProlleSerTyrLeuLysGlySerSerGlyGlyProLeuLeuCysPro 161
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Combined open reading frames of the hepatitis C virus (HCV) cDNA in clones 141 11b, 7f, 7e, 8h, 33c, 40b, 37b, 35, 36, 81, 32, 33b, 25c, 14c,
                  19 GGTAGTGTGGTCATTGTGGGTAGGATCATTTGTCCGGTAGTAGTAGTATCACGGCGTAC 78
                                                                                                                                          61
                                                                     22 AlaGlnGlnThrArgGlyGluGluGlyCysGlnGluThrSerGlnThrGlyArgAspLys 41
                                                                                                                                                                                                          CCCAAGGGTCCTGTCATCCAGATGTATACCAATGTAGACCAAGACCTTGTGGGCTGGCCC
                                                                                                                                                                                                                                                                                                                                                                                                                        ValThrArgHisAlaAspValIleProValArgArgArgGlyAspSerArgGlySerLeu
                                                                                                                                          AsnGlnValGluGlyGluValGlnIleValSerThrAlaAlaGlnThrPheLeuAlaThr
                                                                                                                                                                                                                                                                                                                                                  AlabroGlnGlySerArgSerLeuThrProCysThrCysGlySerSerAspLeuTyrLeu
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ValAsphelleProValGluSerLeuGluThrThrWetArgSer 196
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 559 GTGGACTITATCCCTGTGGAGACCTAGAGACAACCATGAGGTCC 603
GlySerValValIleValGlyArgIleAsnLeuSerGlyAsp--
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Hepatitis C virus; HCV; non-A, non-B hepatitis; NANBH
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Location/Qualifiers 3..5300
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      AAN92097 standard; DNA; 5300 BP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           87US-0122714.
87US-0139886.
88US-0161072.
88US-0263584.
88US-0263584.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              88EP-0310922
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          (updated)
(first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              8f, 33f, 33g and 39c.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                06-MAY-1988;
26-OCT-1988;
14-NOV-1988;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          25-MAR-2003
02-MAR-1990
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            18-NOV-1988;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          31-MAY-1989
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             18-NOV-1987
30-DEC-1987
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  26-FEB-1988
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         EP318216-A
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         AAN92097;
                                                                                                                                                                         139
                                                                                                                                        42
                                                                                                                                                                                                                                                                                                                52
                                                                                                                                                                                                                                                                                                                                                  05
                                                                                                                                                                                                                                                                                                                                                                                   319
                                                                                                                                                                                                                                                                                                                                                                                                                        122
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         142
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  182
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    RESULT 11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        AAN92097
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                      ò
                                                                                                                                                                     g
                                                                                                                                                                                                            à
                                                                                                                                                                                                                                     ద
                                                                                                                                                                                                                                                                      à
                                                                                                                                                                                                                                                                                                              qq
                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                   a
                                                                                                                                                                                                                                                                                                                                                                                                                      ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                     a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 å
                                  qq
                                                                       ò
                                                                                                       å
```

us-09-965-594-14.rnq

```
Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1467
                                                                                                                                                                                          867
                                                                                                                                                                                                                                             34
                                                                                                                                                                                                                                                               287
                                                                                                                                                                                                                                                                                                  1047
                                                                                                                                                                                                                                                                                                                                     1107
                                                                                                                                                                                                                                                                                                                                                                        1167
                                                                                                                                                                                                                                                                                                                                                                                                                                             1287
                                                                                                                                                                                                                                                                                                                                                                                                                                                                154
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1347
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  174
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1407
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      194
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             AAN92103;
                                                                                                                                                                                                                                                                                                                                                       94
                                                                                                                                                                                                                                                                                                                                                                                                                             134
                                                                                                                              Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         RESULT 13
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   AAN92103
                                                                                                    Score:
                                                                                                                                                                                                                                                            qq
                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                      a
                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                             셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  쉽
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    В
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ð
 8888888
                                                                                                                                                                                        g
                                                                                                                                                                                                           ò
                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                       à
                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                          ö
                                                                                 1346
                                                                                                                                                                                           regegegereceretrereceeeeeeeeeeeeeregearntragegeeeerg 1406
                    GCCGGAACGAGGACCATCGCGTCACCCAAGGGTCCTGTCATCCAGATGTATACCAATGTA 1166
                                                                                                            153
                                                                                                                                              SerGlyGlyProLeuLeuCysProAlaGlyHisAlaValGlyIlePheArgAlaAlaVal 173
                                                                        CysGlySerSerAspLeuTyrLeuValThrArgHisAlaAspValIleProValArgArg 133
  6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     The sequence shows the composite cDNA sequence derived from the aligned hepatitis C virus (HCV) cDNA's in clones 141, 11b, 7f, 7e, 8h, 33c, 40b, 37b, 35, 36, 81, 32, 33b, 25c, 14c, 8f, 33f, 33g and 39c. The cDNA
                                                                                                                            1287 CGGGTGATAGCAGGGGCAGCCTGCTGTCGCCCGGCCCATTTCCTACTTGAAAGGCTCC
                                                                                                                                                                                 CysThrArgGlyValAlaLysAlaValAspPheIleProValGluSerLeuGluThrThr
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                prodn. of polynucleotide probes, diagnosis, prevention and treatment
AlaGlyThrArgThrIlcAlaSerProLySGlyProValIlcGlnMetTyrThrAsnVal

        ArgGlyAspSerArgGlySerLeuLeuSerProArgPro1leSerTyrLeuLysGlySer

                                      AspLysAspLeuValGlyTrpProAlaProGlnGlySerArgSerLeuThrProCysThr
                                                                                                                                                        Hepatitis C virus; composite cDNA; probe; vaccine.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Disclosure; Fig. 26; 174pp; English.
                                                                                                                                                                                                                                                                                                                                                                                        Location/Qualifiers
3..5360
/*tag= a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 for
                                                                                                                                                                                                                                                                                                                                     Hepatitis C virus composite probe
                                                                                                                                                                                                                                                                        BP.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               polypeptide(s) and antibodies of infection.
                                                                                                                                                                                                                                                                        AAN90327 standard; cDNA; 5360
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                87US-0132714.
87US-0139886.
88US-0161072.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           88US-0263584
                                                                                                                                                                                                                                                                                                                                                                                                                                                                88GB-0027024
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (CHIR ) CHIRON CORPORATION.
                                                                                                                                                                                                                                                                                                            (updated)
(first entry)
                                                                                                                                                                                                                              ATGAGGTCCCCG 1478
                                                                                                                                                                                                                    MetArgSerPro 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Choo OL,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               WPI; 1989-215054/30.
                                                                                                                                                                                                                                                                                                                                                                        Pan troglodytes
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                18-NOV-1987;
30-DEC-1987;
26-FEB-1988;
26-OCT-1988;
                                                                                                                                                                                                                                                                                                                                                                                                                                                               18-NOV-1988;
                                                                                                                                                                                                                                                                                                           25-MAR-2003
11-NOV-1989
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Houghton M,
                                                                                                                                                                                                                                                                                                                                                                                                                            GB2212511-A
                                                                                                                                                                                                                                                                                                                                                                                                                                             26-JUL-1989
                   1107
                                     46
                                                      1167
                                                                        114
                                                                                          1227
                                                                                                            134
                                                                                                                                              154
                                                                                                                                                              1347
                                                                                                                                                                                                  1407
                                                                                                                                                                                                                                    1467
                                                                                                                                                                                174
                                                                                                                                                                                                                   194
                                                                                                                                                                                                                                                                                         AAN90327
                                                                                                                                                                                                                                                               AAN9032
                                                                                                                                                                                                                                                                          셤
                                    οy
                                                      8
                                                                        ò
                                                                                         셤
                                                                                                            ç
                                                                                                                           셤
                                                                                                                                              ç
                                                                                                                                                              셤
                                                                                                                                                                                 å
                                                                                                                                                                                                 윰
                                                                                                                                                                                                                    ö
                                                                                                                                                                                                                                     셤
ò
```

```
1466
                                                                                                                                                                                                                                                                                                                                                                                                      SerGlyGlyProLeuLeuCysProAlaGlyHisAlaValGlyIlePheArgAlaAlaVal 173
                                                                                                                                                                                                                                                                                              73
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Combined open reading frames of the hepatitis C virus (HCV) cDNAs from
encodes antigens which react with antibodies in patients with non-A non-B hepatitis (NANBH). The CDNA can be used to design probes, or to synthesise polypeptides, which are used to diagnose HCV-induced NANBH, or circles antibodies for immnoassay or treatment, or to produce vaccines. See also AAPP90158, AAN90303-26, and AAN90328-36. (Updated on 25-MAR-2003 to correct PR field.)
                                                                                                                                                                                                                                                                                                                              54 AlaalaGlnThrPheLeuAlaThrCysIleAsnGlyValCysTrpThrValTyrH1sGly
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ArqGlyAspSerArqGlySerLeuLeuSerProArqProlleSerTyrLeuLysGlySer
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ThrSerGlnThrGlyArgAspLysAsnGlnValGluGlyGluValGlnIleValSerThr
                                                                                                   Sequence 5360 BP; 1060 A; 1622 C; 1532 G; 1145 T; 1 other;
                                                                                                                                                                                                                                                                           3 LysLysGlySerValValIleValGlyArgIleAsn-----
                                                                                                                                           5360
175
9
11
                                                                                                                                         Length:
Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                     Gaps:
                                                                                                                                                                                                                                             (1-5360)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    BP.
                                                                                                                                                                                                                                              US-09-965-594-14 (1-197) x AAN90327
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    AAN92103 standard; DNA; 6905
                                                                                                                                         2.97e-72
901.50
90.20%
85.78%
87.35%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            (updated)
(first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ATGAGGTCCCCG 1478
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                MetArgSerPro 197
                                                                                                                                                                       Percent Similarity:
Best Local Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           25-MAR-2003
02-MAR-1990
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    BXANAXAXA
```

1619

153

133

```
Combined open reading frames of the hepatitis C virus (HCV) cDNAs from clones K9-1 through 15e.
                                                                                                                                                                                                                                       1500 TGCGGCTCCTCGGACCTTTACCTGGTCACGAGGCACGCCGATGTCATTCCCGTGCGCGG
                                                                                                                                                                                  114 CysGlySerSerAspLeuTyrLeuValThrArgHisAlaAspVallleProValArgArg
                                                                                                                                                                                                                         154 SerGlyGlyProLeuLeuCysProAlaGlyHisAlaValGlyIlePheArgAlaAlaVal
  74 AlaGlyThrArgThrIleAlaSerProLysGlyProVallleGlnMetTyrThrAsnVal
                                                           AspLysAspLeuValGlyTrpProAlaProGlnGlySerArgSerLeuThrProCysThr
                                                                                                                                                                  134 ArgGlyAspSerArgGlySerLeuLeuSerProArgProIleSerTyrLeuLysGlySer
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Hepatitis C virus; HCV; non-A, non-B hepatitis; NANBH
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 and associated nucleic acids and polypeptide(s)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Claim 3; Figure 47-1 - 47-8; 139pp; English
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Location/Qualifiers 3..7310
                                                                                                                                                                                                                                                                                                                                                                                                                    AAN92106 standard; DNA; 7310 BP.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ΰ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    87US-0122714.
87US-0139886.
88US-0161072.
88US-0191263.
88US-0263584.
88US-0271450.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             88EP-0310922
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Kuo
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Purified hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                   1740 ATGAGGTCCCCG 1751
                                                                                                                                                                                                                                                                                                                                                                                                                                                                           (updated)
(first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            /*tag=
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Roughton M, Choo QL,
                                                                                                                                                                                                                                                                                                                                     MetArqSerPro
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       (CHIR ) CHIRON CORP.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             WPI; 1989-159274/22
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        P-PSDB; AAP92050
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     18-NOV-1987;
30-DEC-1987;
26-FEB-1988;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            18-NOV-1988;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                          25-MAR-2003
02-MAR-1990
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               06-MAY-1988
26-OCT-1988
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           14-NOV-1988
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  31-MAY-1989
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      EP318216-A.
                                                         8
                                                                                                                                                                                                                                                                                                                                     194
                                                                                                                                                                                                                                                                                                                                                                                                                                                AAN92106;
                                                                                                                                                                                                                                                                                                                                                                                            RESULT 14
                                                                                                                                                                                                                                                                                                                                                                                                          AAN92106
                            a
                                                                                                            οχ
                                                                                                                                       g
                                                                                                                                                                  à
                                                                                                                                                                                           g
                                                                                                                                                                                                                         ö
                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                       q
                                                                                                                                                                                                                                                                                                                                    à
                                                                                                                                                                                                                                                                                                                                                               a
                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                   g
                                                                                                                                                                                                                                                                                ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1200 TIGCTGGCGCCCATCACGGCGTACGCCCAGCAGACAAGGGGCCTCCTAGGGTGCATAATC 1259
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1140 CGCAGGGCCGGGAGATACTGCTCGGGCCAGCCGATGGAATGGTCTCCAAGGGGTGGAGG 1199
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    It is a double-stranded nucleotide sequence of the open reading frame (ORF) (tag a) extending through clones 12f to 15e of hepatitis C virus (HCV) cDNa. It can be used to make oligomeric DNa hybridisation probes to detect the presence of HCV nucleic acids in samples. The polypeptide(s) it encodes could be used as immunoassay reagents and vaccines and to generate antibodies useful in diagnosis and passive immunotherapy for HCV infection/non-A, non-B hepatitis.
(Updated on 25-MAR-2003 to correct PR field.)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    33
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ---LeuSerGlyAspThrAlaTyrAlaGlnGlnThrArgGlyGluGluGlyCysGlnGlu
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Sequence 6905 BP; 1421 A; 2082 C; 1946 G; 1456 T; 0 other;
                              Hepatitis C virus; HCV; non-A, non-B hepatitis; NANBH.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              3 LysLysGlySerValVallleValGlyArgIleAsn....
                                                                                                                                                                                                                                                                                                                                                                                                                   Purified hepatitis C virus - and associated nucleic acids and polypeptide(s)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                             Claim 3; Figure 32-1 - 32-7; 139pp; English
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Length:
Matches:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Gaps:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  US-09-965-594-14 (1-197) x AAN92103 (1-6905)
                                                                                  Location/Qualifiers
3..6905
/*tag= a
                                                                                                                                                                                                                                                                                                                                                  Kuo G;
                                                                                                                                                                                                                     B7US-0122714.
B7US-0139886.
88US-0161072.
88US-0191263.
88US-0263584.
88US-0271450.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   4.05e-72
901.50
90.20%
85.78%
87.35%
                                                                                                                                                                                              88EP-0310922
clones 12f through 15e
                                                                                                                                                                                                                                                                                                                                                 Choo QL,
                                                                                                                                                                                                                                                                                                                       (CHIR ) CHIRON CORP
                                                                                                                                                                                                                                                                                                                                                                            WPI; 1989-159274/22
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Percent Similarity:
Best Local Similarity:
                                                        Hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                         P-PSDB; AAP92047
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Alignment Scores:
                                                                                                                                                                                            18-NOV-1988;
                                                                                                                                                                                                                                                                                                                                                 Houghton M,
                                                                                                                                                                                                                                     30-DEC-1987
                                                                                                                                                                                                                     18-NOV-1987
                                                                                                                                                                                                                                                                 06-MAY-1988
                                                                                                                                                                                                                                                                              26-OCT-1988
                                                                                                                                                                                                                                                                                           14-NOV-1988
                                                                                                                                                                  31-MAY-1989
                                                                                                                                       EP318216-A
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    15
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Query Match:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Pred. No.:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Score:
```

à 셤 à 엄 ò g

a à

(updated)

```
19-JUL-2001
             01-NOV-1989
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ..
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Pred.
     a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   2024
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            2145 TCGGGGGGTCCGCTGTTGTTGTTGTTGTTTTAGGCCCGCGGT 2204
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 2025 IGCGGCTCCTCGGACCTTTACCTGGTCACGAGGCACGCCGAIGTCATTCCCGIGCGCCGG 2084
                                                                                                                                                                                                                                                               1665 CGCAGGGGCCGGGAGATACTGCTCGGGCCAGCCGATGGAATGGTCTCCAAGGGGTGGAGG 1724
                                                                                                                                                                                                                                                                                                1725 TTGCTGGCGCCCATCACGGCGTACGCCCAGCACAAGGGGCCCTCCTAGGGTGCATAATC 1784
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   153
                          Ç
                                                                                                                                                                                                                                                                                                                                                                                                                                                             AspLysaspLeuvalGlyTrpProAlaProGlnGlySerArgSerLeuThrProCysThr 113
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CysGlySerSerAspLeuTyrLeuValThrArgHisAlaAspValIleProValArgArg 133
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           154 SerGlyGlyProLeuLeuCysProAlaGlyHisAlaValGlyIlePheArgAlaAlaVal 173
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     174 CysThrArgGlyValAlaLysAlaValAspPheIleProValGluSerLeuGluThrThr 193
                                                                                                                                                                                                                                                                                                                                                                           73
                                                                                                                                                                                                                                                                                                                                                                                                                     93
                                                                                                                                                                                                                                                                                      33
It is a double-stranded nucleotide sequence of the open reading frame (ORF) (tag a) extending through clones K9-1 to 15e of hepatitis C virus (HCV) cDNA. It can be used to make oligomeric DNA hybridisation probes to detect the presence of HCV nucleic acids in samples. The polypeptide(s) it encodes could be used as immunoassay reagents and vaccines and to generate antibodies useful in diagnosis and passive immunotherapy for HCV infection/hon-A, non-B hepatitis.
(Updated on 25-MAR-2003 to correct PR field.)
                                                                                                                                                                                                                                                                                      ---LeuSerGlyAspThrAlaTyrAlaGlnGlnThrArgGlyGluGluGlyCysGlnGlu
                                                                                                                                                                                                                                                                                                                                                                          AlaAlaGlnThrPheLeuAlaThrCysIleAsnGlyValCysTrpThrValTyrHisGly
                                                                                                                                                                                                                                                                                                                                                                                                                   74 AlaGlyThrArgThrIleAlaSerProLysGlyProValIleGlnMetTyrThrAsnVal
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ArgGlyAspSerArgGlySerLeuLeuSerProArgProIleSerTyrLeuLysGlySer
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  GACCAAGACCTTGTGGGCTGGCCCGCTCCGCAAGGTAGCCGCTCATTGACACCCTGCACT
                                                                                                             Sequence 7310 BP; 1491 A; 2217 C; 2058 G; 1540 T; 4 other;
                                                                                                                                                                                                                                            3 LysLysGlySerValValIleValGlyArglleAsn----
                                                                                                                                             Length:
Matches:
Conservative:
                                                                                                                                                                             Mismatches:
Indels:
                                                                                                                                                                                                                      (1-7310)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   BP.
                                                                                                                                                                                                                       US-09-965-594-14 (1-197) x AAN92106
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 AAN90336 standard; DNA; 7310
                                                                                                                                       4.35e-72
901.50
90.208
85.788
87.358
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      2265 ATGAGGTCCCCG 2276
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   MetArgSerPro 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              (updated)
                                                                                                                                                                              Similarity:
                                                                                                                                                                   Percent Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             25-MAR-2003
                                                                                                                                   Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                               1845
                                                                                                                                                                                                                                                                                                                                                                                                                                       1905
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1965
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                194
                                                                                                                                                                                                                                                                                      12
                                                                                                                                                                                                                                                                                                                                                                          54
                                                                                                                                                                                                                                                                                                                                                                                                                                                              46
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        114
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                134
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      2085
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        AAN90336
                                                                                                                                                                                       Query Match:
                                                                                                                                                                              Best Local
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              RESULT 15
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       AAN90336
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SXXXE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      용
  8$888888888
                                                                                                                                                                                                                                            ð
                                                                                                                                                                                                                                                                 a
                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                        윱
                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                    셤
                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                             à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     à
```

```
1844
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   73
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         53
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     93
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      derived by
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          The sequence shows a composite hepatitis C virus (HCV) cDNA, derived by aligning clones k9-1 through 15e in 5'-3' direction. The cDNA encodes antigens which react with antibodies in patients with non-A non-B hepatitis (NANBH). The cDNA can be used to design probes, or to synthesise polypeptides, which are used to diagnose HCV-induced NANBH, to raise antibodies for immunoassay or treatment, or to produce vaccines. See also AAP90288, and AAN90303-35.

(N.B. This record was resubmitted to correct errors in the sequence.) (Updated on 25-MAR-2003 to correct PR field.)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         74 AlaGlyThrArgThrIleAlaSerProLysGlyProVallleGlnMetTyrThrAsnVal
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ---LeuSerGlyAspThrAlaTyrAlaGlnGlnThrArgGlyGluGluGlyCysGlnGlu
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     34 ThrSerGlnThrGlyArgAspLysAsnGlnValGluGlyGluValGlnIleValSerThr
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      54 AlaAlaGlnThrPheLeuAlaThrCysIleAsnGlyValCysTrpThrValTyrHisGly
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Hepatitis C virus gene - used for prodn. of polynucleotide probes, polypeptide(s) and antibodies for diagnosis, prevention and treatment of infection.
                                                                                            ds.
                                                                                       Hepatitis C virus; cDNA; clone 15e; clone k9-1; probe; vaccine;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Sequence 7310 BP; 1495 A; 2218 C; 2058 G; 1539 T; 0 other;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Length:
Matches:
Conservative:
Mismatches:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              3 LysLysGlySerValValIleValGlyArgIleAsn---
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   US-09-965-594-14 (1-197) x AAN90336 (1-7310)
                                         Composite hepatitis C virus (HCV) cDNA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Disclosure; fig 47; 235pp; English.
                                                                                                                                                                                                                                                                                                                         87US-0122714.
87US-0139886.
88US-0161072.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             4.35e-72
901.50
90.20%
85.78%
87.35%
                                                                                                                                                                                                                                                                             88GB-0027024
                                                                                                                                                                                                                                                                                                                                                                                               88US-0263584
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Kuo
                                                                                                                                                                                                                                                                                                                                                                                                                                           (CHIR ) CHIRON CORPORATION.
(first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Houghton M, Choo QL,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     WPI; 1989-215054/30
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Best Local Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             P-PSDB; AAP90288
                                                                                                                                     Pan troglodytes
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Percent Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Alignment Scores:
                                                                                                                                                                                                                                                                             18-NOV-1988;
                                                                                                                                                                                                                                                                                                                                               30-DEC-1987;
26-FEB-1988;
                                                                                                                                                                                                                                                                                                                                                                                               26-0CT-198B;
                                                                                                                                                                                   GB2212511-A
                                                                                                                                                                                                                                                                                                                         18-NOV-1987;
                                                                                                                                                                                                                                 26-JUL-1989
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             13
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Query Match:
```

```
1905 GCCGGAACGAGGACCATCGCGTCACCCAAGGGTCCTGTCATCCAGATGTATACCAATGTA 1964
   1965
ద
   õ
```

Search completed: August 30, 2003, 19:47:57 Job time : 192.939 secs

```
Compugen Ltd.
GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compug
```

OM protein - nucleic search, using frame\_plus\_p2n model

August 30, 2003, 19:20:43; Search time 1910.31 Seconds (without alignments) 2506.388 Million cell updates/sec Run on:

1 MKKKGSVVIVGRINLSGDTA......VAKAVDFIPVESLETTMRSP 197 US-09-965-594-14 1032 Title: Perfect score: Sequence: Scoring table:

BLOSUM62 Xgapop 10.0, Xgapext ( Ygapop 10.0, Ygapext ( Fgapop 6.0, Fgapext ( Delop 6.0, Delext

22781392 seqs, 12152238056 residues Searched:

45562784 Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0 Maximum DB seq length: 200000000

Post-processing: Minimum Match 0% Maximum Match 100% Listing first 45 summaries

Command line parameters:
-WODEL-frame+\_pln.model\_DEV=xlp
-WODEL-frame+\_pln.model\_DEV=xlp
-WODEL-frame+\_pln.model\_DEV=xlp
-U-Cgn2\_luSp70\_spool/US09965594/runat\_29082003\_151919\_28322/app\_query.fasta\_1.2872
-U-Cgn2\_luSp70\_spool/US09965594/runat\_29082003\_151919\_28322/app\_query.fasta\_1.2872
-U-Cgn2\_luSp70\_spool/US09965594-runat\_spool/US08\_SP0\_nusnd0.cdi -LIST-45
-UNTTS-bits -START-1 = NDD--1 -MATRIX-blosum62 -TRANS-humand0.cdi -LIST-45
-UOTRET-ptc -NORM-ext -HEAFSIZE-500 -MINIEN-0 -MAXIEN-200000000
-USER-US0996554\_CGN\_l\_12630\_drunat\_29082003\_151919\_28322 -NCPU-6 -ICPU-3
-NO\_MATP -LIARGOUERY -NEG\_SCORES-0 -MAIT -DSPBLCK=100 -LONGLOG
-DEV\_TIMEOUT-120 -WARN\_TIMEOUT-30 -THERADS-1 -XGAPEXT-0.5 -FGAPOP-6
-FGAPEXT-7 -YGAPOP-10 -YGAPEXT-0.5 -DELOP-6 -DELEXT-7

EST: \* Database :

em\_gss\_inv:\* em\_gss\_hum:\* em\_estfun:\* em\_estba:\*
em\_esthum:\*
em\_estin:\*
em\_estov:\*
em\_estpl:\*
em\_estpl:\*
em\_estro:\*
em\_htc:\*
gb\_estl:\* em\_estom: \* gb\_est2:\* gb\_htc:\* gb\_est4:\*
gb\_est5:\* gb\_est3:\* 10:

em\_gss\_rod:\* em\_gss\_phg:\* em\_gss\_vrl:\* gb\_gssl:\* em\_gss\_vrt:\* em\_gss\_fun:\* em\_gss\_mam:\* em\_gss\_mus:\* em\_gss\_pro:\* em\_gss\_pln:\* 14: 15: 17: 18: 20:

29: gb\_gss2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

	escription	04699	92487 AGENC	01625 BJ001	121 BJ024	16176 BJ016	03316 60186	26101 AGENC	54	43	24	27	23	90	88	24	48	65	74	79	95	84	24	0.7	9	8	18	85	27	66	<u>ت</u>	14	8	42	86	290	872	960	73186	5608	78418	85514	94182	0928211	BF180599 601808704	B763743	
SUMMARIES	Ω		48	62	BJ024121	17	31	2	54	43	24	27	23	90	æ	24	4	65	74	79	05	84	24	5	<b>4</b> 0	80	18	82	27	6	S.	7	8	4.2	88	23	87	98	57318	5560	17841	38551	F79418	092821	BF180599	B76374	
	DB	۱.	~	~	~	~	$\overline{}$	m		$\sim$	0	0	0		m		~	m	~	₹	0	m		0	œ	a	σ	_	-	•	-		m	·	4	m	m	0	13	12	10	10	10	13	2	14	
	ength	. %	•	_	-	10	96	$\sim$	14	~	$\sim$	-	$\sim$	0	m	-		93	æ	0	3	0	66	6	96	14	œ	63	2	n	യ	⊶ .	❤ '	9	oo.	$\sim$	◂	n	93	0	$\sim$	σ.	8	$\circ$	ω.	$\sim$	
ď	Quer		。																					•			•		•		•		•	•	•	•	•	•	•	•	•		•	•	6.8	•	
	Score			66	66	σ	98.5	89	5	97.5	96	Φ	95.5	95	σ	₹.	94.5	÷	σ	ë,	Э.	93.5	έ.	m.	93	93	93	93	σ	ċ	92.5	'n.	ai,	ni.	'n	'n	'n	ď	ä	Ċ.	92	92	92	92		91.5	
	ult No.		7	٣	4	'n	φ	7	œ	o,	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	56	27	28	59	30	<u>ج</u>	32	m m	34	32	36	37	38	39	40	41	42	43	44	45	
	S	. 0		υ			O	ပ				O					O		O			U		ပ	υ	ပ		U	O		ပ							O		U		O		υ	U	O	

## ALIGNMENTS

984 bp mRNA linear EST 21-NOV-2000 601888252F1 NIH\_MGC\_17 Homo sapiens cDNA clone IMAGE:4122276 5', BF304699 Eukaryota, Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. 1 (bases 1 to 984) BF304699.1 GI:11251586 Homo sapiens (human) Homo sapiens EST. BF304699/c LOCUS DEFINITION ORGANISM ACCESSION VERSION KEYWORDS SOURCE RESULT 1

REFERENCE

```
Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                 source
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  BASE COUNT
ORIGIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Pred. No.:
                                        REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
                                                                                                                                                                                                                                                                                                                                                            FEATURES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Score:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Ω
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    d
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               윱
                                                                                                                                                                                                                                                                                         /organism="Homo sapiens"
/mol_type="mRNA"
/Mol_type="mRNA"
/do_type="mRNA"
/do_type="thadGe:412276"
/tissue_type="thadGomyosarcoma"
/tissue_type="thadGomyosarcoma"
/lab_host="hull0B (phage-resistant)"
/lab_host="hull0B (phage-resistant)"
/lab_host="hull0B (phage-resistant)"
/note="Organ: muscle: Vector: poTB7; Site_1: ECORI;
Site_2: XhoI; CDNA made by oligo-dT priming.
Directionally cloned into EcoRY/XhoI sites using the
following 5' adaptor: GGGGCAGG(G). Size-selected >500bp
for average insert size 1: 8kb. Library constructed by
Ling Hong in the laboratory of Gerald M. Rubin (University
of California, Berkeley) using ZAP-CDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies)."
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1199 bp mRNA linear EST 16-AUG-2002
AGENCOURT_8417538 Lupski_sympathetic_trunk Homo sapiens cDNA clone
IMAGE:6192708 5', mRNA sequence.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      119 LeuTyrLeuValThrArgHisAlaAspVallleProValArgArgArgGlyAspSerArg 138
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               100 TrpProAlaProGlnGlySerArgSerLeuThr -- - ProCysThrCysGlySerSerAsp 118
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               -----ACCAGGCACGCAACAACATACATGCAAGGAGACGTGGT---TCCCGC 554
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  154 SerGlyGlyProLeuLeuCysProAlaGlyHisAlaValGlyIlePheArgAlaAlaVal 173
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  493 CGGCACGCTCCGGTCAGGTGCAGC---------TTCCAGCGCCCGGGG 455
                                                                             Email: cgapbs-remail.nih.gov

Tissue Procurement: ArC

CDNA Library Preparation: Ling Hong/Rubin Laboratory

CDNA Library Preparation: Ling Hong/Rubin Laboratory

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Incytte Genomics, Inc.

Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov

Plate: LLCM1005 row: g column: 13

High quality sequence stop: 646.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   NIH-MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished
Contact: Robert Strausberg, Ph.D.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      984
33
33
5
5
22
5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Length:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          US-09-965-594-14 (1-197) x BF304699 (1-984)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      BQ892487
BQ892487.1 GI:22284501
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Homo sapiens (human)
Homo sapiens
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      106.00
45.24%
39.29%
10.27%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     174 CysThrArgGly 177
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1. .984
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Percent Similarity:
Best Local Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               298
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Query Match:
DB:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Pred. No.:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             DEFINITION
AUTHORS
TITLE
JOURNAL
COMMENT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          BASE COUNT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        RESULT 2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         BQ892487
                                                                                                                                                                                                                                                        FEATURES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ð
```

```
//Cissue_type="sympathetic trunk"
//dev_stage="adult, 16 yr"
//dev_stage="adult, 16 yr"
//dev_stage="adult, 16 yr"
//deb.host="bilio#"
//deb.host="bilio#"
//deb.host="bilio#"
//deb.host="bilio#"
//deb.host="bilio#"
//dest_filio#"
//dest_filio#
//dest_fili
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

I (bases it or 1199)

NIH-MCC http://mgc.nci.nlh.gov/.

National Institutes of Health, Mammalian Gene Collection (MGC)

L Onpublished
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-rémail.nih.gov
Tissue Procurement: Dr. James R. Lupski
cDNA Library Preparation: Life Technologies, Inc.
Clone distribution: MCC clone distribution information can be found through the I.M.A.G.E. Consortium/LLML at:
http://image.llnl.gov
Plate: LLAM13555 row: c column: 13
High quality sequence start: 57
High quality sequence start: 57
Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ----CysGlySerSerAsp 118
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        143 -----SerProArgProlleSerTyrLeuLys----------GlySerSerGl 155
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           88 GlnMetTyrThrAsnValAspLysAspLeuValGlyTrpProAlaProGlnGlySerArg 107
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                68 TrpThrValTyrHisGlyAlaGlyThrArgThrIleAlaSerProLysGlyProValIle 87
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           484 TGGGATCCATTTTAATAAAGGGTGCTCTGTTAATCATGGCGCCCACGGGCCCGCTGATA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            574 --- TGTGCTGCCTGCACACACCCCCATGACCGATGTTGGGCTTATGTGGAAACGGCGAG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        119 LeuTyr-LeuValThr------ArgHisAlaAspValIleProValArg----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               28 others
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Matches:
Conservative:
Mismatches;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          211 t
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1. . 1199.
/organism="Homo sapiens"
/organism="Thomosapiens"
/orse="Thomosapiens"
/db.xref="taxon:9606"
/clone="IMAGE:6192708"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Length:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          US-09-965-594-14 (1-197) x BQ892487 (1-1199)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                343 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Fechnologies.
362 c
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            108 SerLeuThrProCysThr
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              /sex-"male"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                10.4
103.50
37.42%
26.45%
10.03%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Percent Similarity:
Best Local Similarity:
Ouery Match:
DB:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ð
```

:::	Oy 112 CysThrCysGlySerSerAspLeuTyrLeuValThrArg 124	Oy 125	Oy 138 ArgGlySerLeuLeuSerProArgProlleSerTyrLeuLysGlySerSer 154   1   1   1   1   1   1   1   1   1	Qy 155 GlyGlyProLeuLeuCysProAlaGlyHisAla 165 	BJ024121 643 bp mRNA	ITION BJ024121 MF01SSA CDNA Oryzias latify MRNA sequence. SION BJ024121 ON BJ024121.1 GI:17377389	_	REFERENCE 1 (bases 1 to 64)  AUTHORS Kohara,Y., Shin-1,T., Kimura,T., Narita,T., Jindo,T. and Takeda,H.	COMMENT CONTACT: Ladasu Shin's Center For Genetice Information National Institute of Genetics	1111 Yata, Mishima, Shizuoka 411-8540, Japan Tel: 81-559-81-685 Fax: 81-559-81-6855	l: ts	/organism="Oryzias latipes" /mol_type="mRNA" /strain="Hd-rR"	/db_xref="taxon:8090" /clone="MF01SSA143112" /sex="mixture of female" /tissue_type="whole embryo"	/dev_stage="segmentation stage 20 - 25" /clone_lib="MF01SSA_CDNA"  BASE COUNT 171 a 148 c 148 g 176 t ORIGIN	Scores:	12.7 99.00 1larity: 33.77% Similarity: 27.81%	Indels: Gaps:	US-09-965-594-14 (1-197) x BJ024121 (1-643)	41 LysAsnGlnValGluGlyGluValGlnIleValSerThrAlaAlaGlnThrPheLeuAla	242 AAAAAIGALGIAGAALCAAAAGALACAALCAALCACALALAGIICIGGIICIACGGGCI	qy 61 ThrCysIleAsnGlyValCysTrpThrValTyrHisGlyAlaGlyThrArgThrIleAla 80
Db 751 AAACCCACCCACCCATGCGCTAAGCCTCCTTTACAAGCCACGGCGGG 810	Oy 155 yGlyProLeuLeuCysProAlaGlyHisAlaValGlyIlePheArgAlaAlaValCysTh 175	<pre>Qy 175 rArgGlyValAlaLysAlaValAspPhelleProValGluSer 189</pre>	RESULT 3 BJ001625/C LOCUS BJ001625 BJ001625 COUNTRY OF THE EST 05-DEC-2001	ON BAJOO1625.1 GI:17364516	KETWORDS EST. SOURCE Oryzias latipes (Japanese medaka) ORGANISM Oryzias latipes Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Buteleostomi;	Actinopterygii; Neopterygii; Teleostei; Buteleostei; Acanthomorpha; Acanthopterygii; Percomorpha; Atherin Beloniformes; Adrianichthyidae; Oryziinae; Oryzias. 1 (bases 1 to 615)	ed of	IIII 1444, MISINGA 5DIZUOKA 411-0340, Japan Tel: 81-559-81-6856 Fax: 81-559-81-6855	source 1. old   /o.gan.sm="Oryzias latipes" /mol_type="mRNA"	/strain="Hd-rR" /db_xref="texes 8090" /clone="MF015SA025C02"	/sex="mixture of female and male" /ilssue_type="whole embryo" //doc etaro="commontation etaro 30 - 25"		Length: Matches:	Percent Similarity: 33.77% Conservative: 9 Best Local Similarity: 27.81% Mismatches: 50 Query Watch: 12 DB: 7	US-09-965-594-14 (1-197) x BJ001625 (1-615)	Oy 41 LysasnginvalglugiygluvalginilevalserthralaalaginthrPheLeuala 60 	Qy 61 ThrCyslleAsnGlyValCysTrpThrValTyrHisGlyAhaGlyThrArgThr11eAla 80	Db 451TGTTGGAGAACCTATCACAGTTCCTGCTTTAGAGCAACGCA 410	SerProLysGlyproVallleGlnMetTyrThrAsnValAspLys	409 GCICIGCGCCAGCAGCICCIGGCCCAGIIGIG	Qy 96 AspLeuValGlyTrpProAlaProGlnGlySerArgSerLeuThrPro 111

US-09-965-294-14	6::1	Db 242 AAAA	Qy 61 ThrC	302	Oy 81 SerP	* ·	Qy 96 AspL Db 380	Qy 112	Db 419 CTGC	Oy 125	Qy 138 ArgG	530	0y 155 GlyG	DP 590 GGAG	RESULT 6	BF20 6018	ACCESSION BF2033	KEYWORDS EST.	ORGANISM HOMO S	Mammai REFERENCE 1 (ba AUTHORS NIH-MG		COMMENT COntac	Tissue	CDNA DNA S CLONE CLONE CLONE CLONE FOUND F	FEATURES SOUTCE		
Db 302TGTTGGAGAACCTATCACAGTTCCTGCTTTAGAGCAACGGCA 343	Torques 10	GCTCTGCGGCGAGGAGCTCCTGGGCCAGTTGTG	Oy 96 AspLeuValGlyTrpProAlaProGlnGlySerArgSerLeuThrPro 111	Db 380ACTCGTGGAGGAGGAGGCGTCACCCGGAGGTGTAGG 418	Oy 112CysThrCysGlySerAspLeuTyrLeuValThrArg 124	בוא כומכשפפפשופרפפשופופמרונומניןיוופפווערורפורפורפונים	Oy 125	Qy 138 ArgGlySerLeuLeuSerProArgProlleSerTyrLeuLysGlySerSer 154	Db 530 AGAGGCCGGACAGGCAGCAGTGGATCTCTGCAGGACGTCTCACGGGGGTTGAATCTCTGCAGGACGTCAGGTGAATCTCTGCAGGACGTGAATCT	Qy 155 GlyGlyProLeuLeuCysProAlaGlyHisAla 165 	RESULT 5	Z	mRNA sequent		×	Eukaryota, Metazoa, Chordata, Craniata, Vertebrata, Euteleostomi, Actinopterygii, Neopterygii, Teleostei; Euteleostei, Neoteleostei;		REFERENCE I (Dases 1 to 734) AUTHORS Kohara,Y., Shin-1,T., Kimura,T., Narita,T., Jindo,T. and Takeda,H. TITE MANAWA EST Drotact in Takeda's lah		Center For Genetic Resource Information National Institute of Genetics 1111 Yata, Mishima, Shipuoka 411-840, Japan	Tel: 81-559-81-6856 Fax: 81-559-81-6855		source 1754 /organism="Oryzias latipes"	/mol_type="mRNA" /strain="Hd-rR" /db_xref="taxon:8090" /clone="MF01SSA025C02" /sex="mixture of female and male"	/tissue_type="whole embryo" //dev_trage="seqmentation stage 20 ~ 25" /clone lib="MF01SSA cDNA"	BASE COUNT 194 a' 181 c 181 g 198 t ORIGIN	Alignment Scores: 15.6 Length: 754 Score: 99.00 Matches: 42 Percent Similarity: 33.77% Conservative: 9 Best Local Similarity: 27.81% Mismatches: 50 Query Match: 12 Gaps: 7

```
316 961 bp mRNA linear EST 06-NOV-2000 5914Fl NIH_MGC_17 Homo sapiens cDNA clone IMAGE:4098578 5',
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         nubblished

ontact: Robert Strausberg, Ph.D.

mail: cgapbs-rémail.nih.gov
issue Procurement: ArC.

CDN Library Preparation: Ling Hong/Rubin Laboratory

cDNA Library Preparation: Ling Hong/Rubin Laboratory

CDNA Library Preparation: Ling Hong/Rubin Laboratory

DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)

Clone distribution: MGC clone distribution information can be

clone distribution can 
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    TCCTCACCTGACCTTCCACATCCAGGTGTCGCCAGCGCTGGTCTGACGGGTGATGGG 529
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ||||
|GGCCGGACAGGCAGCAGTCGGGGGTGAATCTCTGCAGGACGTCTTCACGGCGGGATCA 589
                                                                                                                                                                                          ------TGTTGGAGAACCTATCACAGTTCCTGCTTTAGAGCAACGGCA 343
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CTGCGGCGGCGAGGAGCTCCTGGGCCAGTTGTG-----379
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     LeuValGlyTrpProAlaProGlnGlySerArgSerLeuThrPro----- 111
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           -------ACTCGTGGAGGACGAAGAGCGTCACCCCGGAGCTGTAGG 418
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 -----CysThrCysGlySerSerAspLeuTyrLeuValThrArg------ 124
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        GlySerLeuLeuSerProArg------ProIleSerTyrLeuLysGlySerSer 154
                                                                                                           AsnGlnValGluGlyGluValGlnIleValSerThrAlaAlaGlnThrPheLeuAla 60
                                                                                                                                                                                                                                                                                                                                  CysileAsnGlyValCysTrpThrValTyrHisGlyAlaGlyThrArgThrlleAla 80
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ProLys------GlyProVallleGlnMetTyrThrAsnValAspLys 95
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      yota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
lia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            nases 1 to 961)
MGC http://mgc.nci.nih.gov/.
nal Institutes of Health, Mammalian Gene Collection (MGC)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 /clone="IMAGE:4098578"
/tissue_type="rhabdomyosarcoma"
/lab_host="PHIOB (phage-resistant)"
/clone_lib="NIH_MGC_17"
/note="Organ: muscle; Vector: poTB7; Site_1: EcoR1;
Site_2: Kho1; cDNA made by oligo-dr priming.
Directionally cloned into EcoR1/XhoI sites using the
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      GlyProLeuLeuCysProAlaGlyHisAla 165
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1. 961
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
(1-197) x BJ016176 (1-754)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    316.1 GI:11096902
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                sapiens (human)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    sequence.
```

Ŋ

Score:

a

ò

g

à

a ò a

ò

```
297
                                                                                                                                                Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                           45
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           65
                                                                                                                                                                                                                              Query Match:
DB:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    VERSION
KEYWORDS
SOURCE
ORGANISM
                                                                                                                                                               ..
Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       LOCUS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        AUTHORS
TITLE
JOURNAL
MEDLINE
                                                                                               BASE COUNT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       RESULT 8
AK080545
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ACCESSION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        REFERENCE
                                                                                                                                                              Preď.
                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                           요
                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                         셤
                                                                                                                                                                                                                                                                                                                                                                                                                                           õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         bUyzblUl 1403 bp mRNA linear EST 20-AUG-2002
AGENCOURT_8752655 NIH_MGC_130 Mus musculus cDNA clone IMAGE:6335718
5′, mRNA sequence.
             for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies)."
                                                                                                                                                                                                                                                                                                                                                                                                                           160
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          160 ysProAlaGlyHisAlaValGlyIlePheArgAlaAlaValCysThrArgGlyValAlaL 180
                                                                                                                                                                                                                                                                                             101 ProAlaProGlnGlySerArgSerLeuThrProCysThrCysGlySerSer-AspLeuTy 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Mus musculus Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mamalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 1403) NIH-MGC http://mgc.nci.nih.gov/.
                                                                                                                                                                                                                                                                                                                                                            120 rLeuValThrArgHisAlaAspValIleProValArgArgArgGlyAspSerArgGlySe
                                                                                                                                                                                                                                                                                                                                                                                                                          140 rLeuLeuSer-ProArgProlleSerTyrLeuLysGlySerSerGlyGlyProLeuLeuC
                                                                                                                                                                                                                                                                                                                          642 ccgcrccrccrccrccrcggcA-----TGAGAATGTGGAGCTCAGAGACATGCT
                                                                                                                                                                                                                                                                                                                                                                                           591 GACTCTGAGTGCAGGAGCGGACAGAGATGCGGATCGGAAGCGCGGGGATGCTGATGGGGG
                                                                                                                                                                                                                                                                                                                                                                                                                                            477 GTCCTGCAGTGCTGCTGTAGGTGCTGTCCGTCAGCTTTTTGTGATGGTGGTGACGCTG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ğ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CDNA Library Preparation: ResGen, Invitrogen Corp
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Apencourt Bioscience Corporation
Clone distribution: MGC clone distribution information
found through the I.M.A.G.E. Consortium/LLNL at:
http://imagellnl.gov
Plate: LLAM13798 row: j column: 07
High quality sequence stop: 101.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Ph.D.
Ph.D. and Nancy IPh.D.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 /organism-"Mus musculus"
/mol_type-"mRNA"
/db_txref-"taxon:10090"
/clone-"lMAGE:6335718"
/lab_host-"DH10B (phage-resistant)"
                                                                                                                                              961
32
11
32
7
following 5' adaptor: GGCACGAG(G)
                                                                                                                                             Length:
Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                Gaps:
                                                                                                                                                                                                                                                             US-09-965-594-14 (1-197) x BF203316 (1-961)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Mus musculus (house mouse)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        BQ926101
BQ926101.1 GI:22341132
                                                                                                                                             23.8
98.50
52.44%
39.02%
9.54%
                                                                                                                                                                             Percent Similarity:
Best Local Similarity:
Query Match:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         180 ys 180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        417 AA 416
                                                                                                                               Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          EST.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       source
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      VERSION
KEYWORDS
SOURCE
ORGANISM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        B0926101/c
LOCUS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      DEFINITION
                                                                                                                                                Pred. No.:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     AUTHORS
TITLE
JOURNAL
COMMENT
                                                                                 BASE COUNT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ACCESSION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       REFERENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        FEATURES
```

RESULT

```
AKO80545

AKO80545

Mus musculus 7 days neonate cerebellum cDNA, RIKEN full length enriched library, clone:A730082L10 product:weakly similar to zinc finger protein (fragment) [Mus musculus], full insert sequence.

AKO80545

AKO80545.

AKO80545.

AKO80545.

Mus musculus (house mouse)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1087 GGCACGGTGATGTGGTTACTACGCGGCAGACCGGCTCCAACGGCGCTCTTAACAGACGC 1028
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1321 TGTCGGGGGCGCGTTGGCGCATACCCCGGGTCGAGGTCAGCGCCGCTTGTATACA 1262
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1249 GGGGTA---TGGTTATCACGGGCTGGGGCAGGTACT-------TCCCCTAAAAGCG 1205
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1144 AA---GAGAGGCACCTGACGCCCCTCCCCTGTGGGCNTGTCGATAATAACAAATGTGCAG 1088
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1378 GGGTGTGTTCANCGGTCAGGACAGGTCGCC---GCACACTCGACGCGTCGGCCAGAGCT 1322
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                -----CAG 1250
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      84 lyProValIleGlnMetTyrThrAsnValAspLySAspLeuValGlyTrpProAlaProG 104
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              137 ----SerArgGlySer-----LeuLeuSerProArgProlleSerTyrLeuLysG 152
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                84
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Musinae; Mus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     GluGlyGluValGlnIleValSerThrAlaAlaGlnThrPheLeuAlaThrCysIleAsn 64
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             11 GlyArglleAsnLeuSerGlyAspThrAlaTyrAlaGlnGlnThrArgGlyGluGluGly 30
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   104 InGlySerArgSerLeuThrProCysThrCysGlySerSerAspLeuTyrLeuValThrA
/clone_lib="NIH_MGC_130"
/note="Organ: otcoysts; Vector: pCMV-SPORT6.1.ccdb;
/note="Organ: otcoysts; Vector: pCMV-SPORT6.1.ccdb;
Priner: Oligo dT. Average insert size 1.95 kb.
Constructed by ResGen, Invitrogen Corp. Note: this is
NIH_MGC Library.
521 c 237 g 345 t 3 others
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        GlyValCysTrpThrValTyrHisGlyAlaGlyThrArgThrIleAlaSerProLys--G
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                124 rgHisAlaAspValIleProValArgArgGlyAsp------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        31 CysGlnGluThr ----- SerGlnThrGlyArgAspLysAsnGln-----
                                                                                                                                                                                                                                                                                  1403
49
12
62
43
                                                                                                                                                                                                                                                                                     Length:
Matches:
Conservative:
Mismatches:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Carninci, P. and Hayashizaki, Y.
High-efficiency full-length cDNA cloning
Meth. Enzymol. 303, 19-44 (1999)
99279253
                                                                                                                                                                                                                                                                                                                                                                                                      Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        US-09-965-594-14 (1-197) x BQ926101 (1-1403)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              152 lySerSerGlyGly 156
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   GAGAAGGCGGCGGT 954
                                                                                                                                                                                                                                                                                     38.5
98.50
36.75
29.52
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1261 GAGGGGAAA----
                                                                                                                                                                                                                                                                                                                                             Percent Similarity:
Best Local Similarity:
```

```
/translation="bsclpaaspgrslltprgdgfflkeklsaaravgpgpsyargvs
Yrkokaadpgrcqrkvcaarrsbsclltbrrrrrchypygphtyglygppage
AqaacrapqyPthprephgryvppgcaagllpalaarqvpvgpyrgregprraprhs
pkpyptalgfsprgdggpappllapangrsvglll"
                                                                                                                                                                                                                                                                                                        /db_xref="FANTOM_DB:A730082L10"
Adb_xref="taxon:10090"
/clone="4730082L10"
/tissue_type="cerebellum"
/tissue_type="cerebellum"
/dev_stage="7 days neonate"
<1. 587
cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      86 ValileGinMetTyrThrAsnValAspLysAspLeuValGlyTrpProAlaProGlnGly 105
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ------CCTCCTCAAGGA 389
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  126 AlaAspVallleProValArg---ArgArgGlyAspSerArgGlySerLeuLeuSerPro 144
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             145 ArgProlle------SerTyrLeuLysGlySerSerGlyGlyProLeu 158
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              486 AAGCCGGTTCCTACAGCCTTGGGTTTCTCGTTTGGGCAGGGTGGGCCTGCTCCTCCCCTC 545
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            -----CAAGIGCCGCATCCGCCIGGG--------CGGCCACAIGGAACA 371
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     106 SerArgSerLeuThrProCysThrCysGlySerSerAspLeuTyrLeuValThrArgHis 125
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    LeuCysProAla -- GlyHisAlaValGlyIlePheArgAlaAlaValCysThrArgGly 177
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   46 GlyGluValGlnIleValSerThrAlaAlaGlnThrPheLeuAlaThrCyslleAsnGly 65
                                                                         Division of Experimental Animal Research in Riken contributed to
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              66 ValCysTrpThrValTyrHisGlyAlaGlyThrArgThrIleAlaSerProLysGlyPro
                                                                                                                                                                                                                                                                                                                                                                                                                                                                         /note="unnamed protein product; putative weakly similar to zinc finger protein (fragment) musculus) (PIR 1148722, evidence: FASTY, 50.7%1D, 57.6%length, match=601)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1141
46
11
52
50
                                                                                               prepare mouse tissues.
Please visit our web site for further details.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Matches:
Conservative:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               303 GGAGAAGCGCAGGCAGCTGGACGCGCGCCGCAG----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Mismatches:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                228
                                                                                                                                         URL:http://genome.gsc.riken.go.jp/
URL:http://fantom.gsc.riken.go.jp/
Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            /protein_id="BAC37940.1"
/db_xref="GI:26348601"
                                                                                                                                                                                                                                   /organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Gaps:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       US-09-965-594-14 (1-197) x AK080545 (1-1141)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                353
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         /note-"putative"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       /note-"putative"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       372 GTGGTG------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                316 c
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        33
98.00
35.85%
28.93%
9.50%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                244 a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Percent Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     polyA_signal
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 polyA_site
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               159
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          336
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Query Match:
                                                                                                                                                                                                                    source
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Best Local
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Pred. No.:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             BASE COUNT
                                                                                                                                                                                            FEATURES
    COMMENT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ORIGIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 d
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Kawai,J., Shinagawa,A., Shibata,K., Yoshino,M., Itoh,M., Ishii,Y.,
Arakawa,T., Hara,A., Fukunishi,Y., Konno,H., Adachi,J., Fukuda,S.,
Arawa,K., Izawa,M., Nishi,K., Kiyosawa,H., Kondo,S., Yamanaka,I.,
Saito,T., Okazaki,Y., Gojobori,T., Bono,H., Kasukawa,T., Saito,R.,
Kadota,K., Matsuda,H., Ashburner,M., Batalov,S., Casavant,T.,
Fleischman,W., Casterland,T., Gissi,C., King,B., Kochiwa,H.,
Kuchl,P., Lewis,S., Matsuo,Y., Nikaido,I., Fesole,G., Tomita,M.,
Wagner,L., Washio,T., Sakai,K., Okido,T., Furuno,M., Aono,H.,
Baldarelli,R., Barsh,G., Blake,J., Boffelli,D., Bojunga,N.,
Carninci,P., de Bonaldo,M., Brownstein,M.J., Bult,C.,
Fletcher,C., Fujita,M., Gariboldi,M., Gustincich,S., Hill,D.,
Marchionni,L., Mashim,J., Mazzacelli,J., Mombaerts,P., Nordone,P.,
Ring,B., Ringwald,M., Rodriguez,I., Sakamoto,N., Sasaki,H.,
Toyo-oka,K., Wang,K.H., Weitz,C., Whittaker,C., Wilming,L.,
Wynshaw-Boris,A., Yoshida,K., Hasegawa,Y., Kawaji,H., Kohtsuki,S.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Kitaunai, T., Tashiro, H., Itoh, M., Sumi, N., Ishi, Y., Wakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwake, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watshiki, M., Yonda, Y., Ishikawa, T., Ozawa, K., Tamaka, T., Marsuura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y., Riken integrated sequence analysis (RISA) system--384-format Genome Res. 10 (11), 1757-1771 (2000)
       Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y. Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes Genome Res. 10 (10), 1617-1630 (2000)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Sanagawa 230-0045, Japan (E-mail:genome-reségsc.riken.go.jp, URL:http://genome.gsc.riken.go.jp/, Tel:81-45-503-9212, Pax:81-45-503-9216)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Submitted (16-APR-2002) Yoshihide Hayashizaki, The Institute of
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Functional annotation of a full length mouse cDNA collection Nature 409 (6821), 685-690 (2001)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Muramatsu, M. and Hayashizaki, Y.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Direct Submission
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 11217851
10349636
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               11076861
                                                                                                                                                            MEDLINE
PUBMED
REFERENCE
AUTHORS
                                                                                                                                                                                                                                                                                                                                                                                                                                              JOURNAL
MEDLINE
PUBMED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   REFERENCE
AUTHORS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    MEDLINE
PUBMED
REFERENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         REFERENCE
AUTHORS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   JOURNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       TITLE
JOURNAL
PUBMED
                                                AUTHORS
                                                                                                                                           JOURNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        AUTHORS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       TITLE
                                                                                            TITLE
                                                                                                                                                                                                                                                                                                                                                                                                     TITLE
```

see Closc TJ, Wing R, Kleinhofs A, Wise R (2001) Genetically and physically anchored EST resources for barley genomics. Barley Genetics Newsletter 31:29-30. (http://wheat.pw.usda.gov/ggpages/bgn/31/cover.html)* ORIGIN	: 22.7 97.50 97.83 39.87 arity: 30.72 10	US-09-965-594-14 (1-197) x BF631437 (1-779)  Oy 49 GlnIleValSerThrAlaAlaGlnThrPheLeuAlaThrCysIleAsnGlyValCySTrp 68	88	Db 169 TCGGCGGGTGGGACCCGGAGTCGTGGGCGACTCGGTGTACGTGTACGACTTCT 228  Qy 118 -AspLeuTyrLeuValThrArgHishlaAspVallleProValArgAr 133	QY         153 rSerGly-GlyProLeuLeuCysProAla	BF863244 BF863244 BF863244 DCCUS DEFINITION FOR ACCESSION FOR FOR FOR FOR ACCESSION FOR FOR FOR ACCESSION FOR FOR FOR FOR ACCESSION FOR FOR FOR FOR ACCESSION FOR FOR FOR FOR FOR ACCESSION FOR
Db 546 CTAGCCCCAGCGAATGGGAGGTCTGTAGGGTTGGCCCTGTAAGTGTGAGTGGGGA 602  OY 178 VALALALYSALAVALASPPHELIEPTOVALGIUSErLeuGluThrThrWet 194  11:	RESULT 9 BF631437 LOCUS BF631437 BF631437  BF631437  BF631437  BF631437  BF631437  BF631437  BF631437  BF631437  BF631437  BF631437  BF631437  BF631437  BF631437  BF631437  BF631437  BF631437.2 GI:13092107	KEYWORDS EST.  SOURCE Hordeum vulgare subsp. vulgare ORGANISM Hordeum vulgare subsp. vulgare ORGANISM Hordeum vulgare subsp. vulgare Spermatophyra; Lillopsida; Poales; Poaceae; Pooideae : Triticeae; Hordeum. REPERENCE 1 (bases 1 to 779) AUTHORS Wing, M. Close, T. J. Kleinhofs, A., Wise, R., Begum, D., Frisch, D., Yulgare, M., Panny, D., Palner, M., Rambo, T., Simmons, J., Choi, D., Benton	'R.D., Oates, R. and Main, D.  TILE Development of a genetically and physically anchored EST resource for barley genomics: Morex drought-stressed seedling shoot cDNA library JOURNAL Unpublished COMMENT On Dec 19, 2000 this sequence version replaced gi:11895595,	Clemson University Genomics Institute Clemson University 100 Jordan Hall, Clemson, SC 29634, USA 101 Jordan Hall, Clemson, SC 29634, USA Tel: 864 656 4293 Email: rwing@clemson.edu Total hg bases = 455 Seq primer: AATTAAACCCTCACTAAAGGG High quality sequence stop: 588. Location/Qualifiers	<pre>gource 1779 /organism-"Hordeum vulgare subsp. vulgare" /mol_type="mRNA" /cultivar="Morex" /db_xref="taxon:112509" /clone="HYPNED6015P05f" /tissue_type="Seedling shoot" /lab_host="TJC121" /clone_lib="Hordeum vulgare seedling shoot EST library</pre>	HVCDMANDUZ (Denydration stress)*  Anote—"Vector: lambda2AP; Site_]: ECCR]; Site_2: Xhol; Seeds were surface sterilized then germinated under axenic conditions in the dark at room temperature on filter paper with water, nystatin and cefotaxime in covered crystallization dishes. Five-day old seedlings were incubated at 90% RH for 24 hr. Shocts were then harvested, total RNA was prepared, poly(A) RNA was purified, one primary unamplified cDNA library was made, 600000 pfu were in vivo excitsed to give pBluescript SK(-) cDNA phagemids. These steps were performed in the TJ Close laboratory at the University of California, Riverside (Choi, Close, Fenton). Phagemids were plated and plicked at the Clemson University Genomics Institute (UGIS) (Begum, Palmer, Frisch, Atkins and Wing). Plasmid DNA preparations, DNA sequencing and sequence analysis were performed at CUGI (Wing, Yu, Frisch, Henry, Simmons, Oates, Rambo, Main). The sequence has been trimmed to remove vector sequence and contains a minimum of 100 bases of phred value 20 or above. For more details on library preparation and sequence analysis see http://www.genome.clemson.edu/projects/barley. To order this clone see http://www.genome.clemson.edu/projects/barley. To order

α

```
Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            176
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Percent Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Query Match:
                                                                                                                                                                                                                                                                                                                                                                                                                    source
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            BASE COUNT
ORIGIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ..
Q
  AUTHORS
TITLE
JOURNAL
COMMENT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    RESULT 12
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                BF307233
                                                                                                                                                                                                                                                                                                                                                                                       FEATURES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   엄
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                     /clone_lib="C. reinhardtii CC-1690, Stress condition I, normalized, Lambda Lap II
//clone_lib="Vector: pBluescript II SK-; Site_l: ECORI; Site_2:
//core="Vector: pBluescript II SK-; Site_l: ECORI; Site_2:
//core: This library, constructed by John Davies and Jeffrey
//complete Constructed by John Davies and Jeffrey
//constructed by John Davies and Jeffrey
//complete Constructed by John Davies and Jeffrey
//constructed by Superinfection with Exassist (Stratagene)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              BF182274 846 bp mRNA linear EST 31-OCT-2000 601804028F1 NCI_CGAP_Mam5 Mus musculus cDNA clone IMAGE:4035102 5',
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     171 CACCACCATACCCTTGCTCTCACGTTGCTCACACAAATTATGCCCATACGGGCCACTA 230
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      107 ArgSerLeuThrProCysThrCysGlySerSerAspLeuTyrLeuValThrArgHisAla 126
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  127 AspValIle........ProValArgArgArgGlyAspSerArg-.... 138
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       87 IleGlnMetTyrThrAsnValAspLysAspLeuValGlyTrpProAlaProGlnGlySer 106
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                231 ACAAAGTIACATACACGG-----AAGGACCAGCGCGCTIGGCCACCCCCTIGGAGCCG 284
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               345 GAGATCATTIGCATGIGGCTTTAGTCACCCCAAGAAGAGCCTGGGAGTGGGCATTTAIAA 404
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          405 GAAGGGGACGGGAATTCCGTTTGCGGAAAAGTGAGCGCCCCAAGGTCTGACCAAGTGCTA 464
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalla; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 846)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  71 TyrHisGlyAlaGlyThrArgThrIleAlaSerProLys------GlyProVal 86
                                                                                                        /organism="Chlamydomonas reinhardtii"
/mol_type="mrRNA"
/strain="CC-1690 wild type mt+ 21gr"
/db_xref="taxon:3055"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  151 LysGlySerSerGlyGlyProLeuLeuCysProAlaGly 163
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   165 CTCCAAGGCAGCAATGGGAGCCTTTCGCGGGTGTGCGGGT 503
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          701
32
14
14
22
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Length:
Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               US-09-965-594-14 (1-197) x BF863244 (1-701)
                                                     Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Mus musculus (house mouse)
Mus musculus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    mRNA sequence.
BF182274
BF182274.1 GI:11060416
                            Email: chauser@duke.edu
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        27.6
96.00
40.71%
28.32%
9.30%
Fax: 919 613 8177
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Percent Similarity:
Best Local Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               173
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Aliqnment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Query Match:
                                                                                    source
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              LOCUS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            BASE COUNT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Pred. No.:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  BF182274/c
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ORGANISM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          REFERENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ACCESSION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            VERSION
KEYWORDS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        RESULT 11
                                                        FEATURES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            SOURCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
```

```
/crganism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/strain="C57BL/6J"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="IMAGE:405102"
/tissue_type="tumor, gross tissue"
/dev_stage="T months"
/dev_stage="T months"
/lab_host="DH10B"
/clone=lib="NCI_CGAP_Mam5"
/clone=lib="NCI_CGAP_Ma
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            709 GGTICTTCAGTCCCAAGTGGGCTGGAGAAGTGAACACAGAGAAAGAGAGGGCGGTCCTCA 650
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    104 nGlySerArg-----SerLeuThrProCysThrCysGlySerSerAspLeuTyrLeuVa 122
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              556
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 142 uSerProArgProIleSerTyrLeuLysGlySerSerGlyGlyProLeu---LeuCysPr 161
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            -----CATCCTTC 710
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             122 lThrArgHisAlaAspValIleProValArgArgArgGlyAspSerArgGlySerLeuLe 142
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   161 oAlaGlyHis---AlaValGlyIlePheArg-------AlaAlaValCySThrAr 176
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    73 GlyAlaGlyThrArg-ThrIleAlaSerProLysGlyProValIleGlnMetTyrThrAs 92
                                                                                                                                                                                                                                                    Email: cgapbs-remail.nih.gov
Tissue Procurement: Lothar Hennisphausen Ph.D., Robin Humphreys
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Robinson Robins
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     NIH-MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1 others
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     757 GGTTCTGCTACCAGAAACAGCCTGGATGAAGAAAGGACCA----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                846
13
36
31
91
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Length:
Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           210 t
                                                                                                                                 Unpublished
Contact: Robert Strausberg, Ph.D.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Gaps:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         US-09-965-594-14 (1-197) x BF182274 (1-846)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           241 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       218 c
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    438 TGGGATGACTGAG 426
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     176 gGlyValAlaLys 180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         35.1
96.00
47.20%
36.80%
9.30%
```

```
Fahrenkrug.S.C., Smith,T.P.L., Freking,B.A., Cho,J., White,J., Vallet,J., Wase,T., Rohrer,G.A., Pertea,G., Sultana,R., Quackenbush,J., and Keele,J.W.

Dordine gene discovery by normalized cDNA-library sequencing and EST cluster assembly
Mamm. Genome 13 (8), 475-478 (2002)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Email: smith@email.marc.usda.gov
Single pass sequencing. Bases called and alt_trimmed with phred
v0.980904.e. Vector identified by cross_match with the -minscore 18
and -minmatch 12 options
                                                                                                                                                              EST 09-JUL-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          61 CAAACCCCGGGCAGCTCTTCGGCCGCATCCCTATGCGCTTCTCGGTGCTGATGCAGATGC 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         121 GTTTCGACGGCCTGCTGGGCTTTCCCGGGGCTTCGTGGACCGGCGCTTCTGGTCATTGG 180
  --- CGGGCATCTCTCCAGGGCCGCTGGACTG 810
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 43 GlnValGluGlyGluValGlnIleValSerThrAlaAlaAlaGlnThrPheLeuAlaThrCys 62
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ----TyrHis-G1 73
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  73 yAlaGlyThrArgThrIleAlaSerProLysGlyProVallleGlnMetTyrThrAsnVa 93
                                                                                                                                                                                                                                                                                                        Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
1 (bases 1 to 407)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               /tissue_type="pooled"
/lab_host="DH10B"
/clone_lib="WARC iPIG"
/note="Vector: pCMV SPORT6; Site_1: Not1; Site_2: Sal1;
Library made from pooled tissue from day 11, 13, 15, 20, and 30 embryos. 145 g 76 t
                                        160 sProAlaGlyHisAlaValGly-----IlePheArgAlaAlaValCys 174
                                                              AW785806 407 bp mRNA linear 117260 MARC IPIG Sus scrofa cDNA 5', mRNA sequence. AW785806
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   63 IleAsnGlyValCysTrpThrVal----------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           407
41
3
52
52
45
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      USDA, ARS, US Meat Animal Research Center
PO Box 166, Clay Center, NE 68933-0166, USA
TTEL: 402 762 4360
Fax: 402 762 4390
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Length:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       /organism="Sus scrofa"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        US-09-965-594-14 (1-197) x AW785806 (1-407)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          FORWARD: AGGAAACAGCTATGACCAT
BACKWARD: GTTTTCCAGTCAGGACG
Plate: 37 row: D column: 16
Seq primer: ATTTAGGTGACACTATAG.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        /mol_type="mRNA"
/db_xref="taxon:9823"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Location/Qualifiers
767 CCTTTTGGGTCCTGA-----
                                                                                                                                                                                                                        AW785806.1 GI:7842582
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Contact: Smith TPL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1. .407
                                                                                                                                                                                                                                                                 Sus scrofa (pig)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         PCR PRimers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Percent Similarity:
Best Local Similarity:
                                                                                                                                                                                                                                                                                         Sus scrofa
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 12226715
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         55
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Query Match:
                                                                                                                                                                                DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       BASE COUNT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Pred. No.:
                                                                                                                                                                                                                                                                                         ORGANISM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              MEDLINE
PUBMED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Alignment
                                                                                                                      RESULT 13
AW785806
                                                                                                                                                                                                                                                                                                                                                   REFERENCE
AUTHORS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                           JOURNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                FEATURES
                                                                                                                                                                                                                                                                                                                                                                                                                                  TITLE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       COMMENT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ORIGIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Score:
                                                                                                                                                                   LOCUS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   윱
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                 g
                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1. 901
/ Organism="Mono sapiens"
/ Mol_type="MRNA"
/ Mol_type="MRNA"
/ Mol_type="MRNA"
/ Mol_type="MRNA"
/ Mol_type="MRNA"
/ Mol_type="MRNA"
/ Mole="Inabdomyosarcoma"
/ Lab_host="Dall08 (phage=resistant)"
/ Clone_llb="NIH_MGC_l7"
/ Note="Organ: muscle; vector: poTB7; Site_l: EcoRI;
/ Anot: Organ: muscle; vector: poTB7; Site_l: EcoRI;
/ Site_2: XhoI: CDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the following 5 adaptor: aGCACGAGGG(0: Size=selected >SOUBp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of Catalagene) and Superscript II RT (Life Technologies)."
44 a 267 c 329 g 161 t
                                                                                                                                                        Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Eukaryota; Eutheria; Primates; Catarrhini; Hominidae; Homo.
El (bases 1 to 901)
In Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-rémail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LINL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC Clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
Plate: LLCM/Ad4 row; c column: 02
High quality sequence start: 6
High quality sequence stop: 684.
BF307233 901 bp mRNA linear EST 21-NOV-2000 601891502F1 NIH_MGC_17 Homo sapiens cDNA clone IMAGE:4137145 5',
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               100 pProAlaProGlnGlySerArgSerLeuThrProCysThrCysGlySerSerAspLeuTy 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      731
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             120 rLeuvalThrArgHisAlaAspValIleProValArgArgArgGlyAspSerArgGlySe 140
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     80 aSerProLysGlyProvalIleGlnMetTyrThrAsnValAspLysAspLeuValGlyTr 100
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   66 ValCysTrpThr-ValTyrHis------GlyAlaGlyThrArgThrIleAl 80
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      716 GCCGCCCCATCCGGG------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      901
37
28
45
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Length:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  US-09-965-594-14 (1-197) x BF307233 (1-901)
                                                                                 BF307233.1 GI:11254342
                                                                                                                          Homo sapiens (human)
Homo sapiens
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    42.5
95.50
38.46%
31.62%
9.25%
                                        mRNA sequence.
BF307233
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Best Local Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         144
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Score:
Percent Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Alignment Scores:
Pred. No.:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Query Match:
DB:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           source
LOCUS
DEFINITION
                                                         ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       BASE COUNT
ORIGIN
                                                                                                                                                                                                       REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    FEATURES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
```

Alignment Scores:  Pred. No.:  95.00  Pred. No.:  95.00  Matches:  60.03  Percent Similarity:  40.13  Conservative:  13.14  Mismatches:  52  Query Match:  13.18  Conservative:  14.00  Conservative:  14.00  Conservative:  14.01  Conservative:  15.18  Conservative:  16.18  Conservative:  17.18  Conservative:  18.18  Conservative:  18.18  Conservative:  18.18  Conservative:	943 TUTOIS	Email: genomics@hri.co.jp HRI human cDNA project; 5'- & 3'-end one pass sequencing: Hellx Research Institute; cDNA library construction: Department of Virology, Institute of Medical Science, University of Tokyo, and Helix Research Institute.
1   1   1   1   1   1   1   1   1   1	RESULT 14 BOG78887 BOG788887 BOG78887 BOG78887 BOG78887 BOG788887 BOG78887 BO	Integraty, non-amplified Library constructed by Life Technologies and donated by J. Lupski, M.D./Ph.D. (Baylor College of Medicine) and is available through Life Technologies."  BASE COUNT 184 a 308 c 253 g 183 t 3 others

```
/organism="Homo sapiens"
//organism="Homo sapiens"
//db_xref="taxon:9606"
//db_xref="taxon:9606"
/clone="WIZPP2002160"
/cell_line="NT2"
//cell_line="NT2"
//orde="Vector: PME18SFL3; mRNA from NT2 neuronal precursor cells after 2. "Weeks refinoic acid (RA) induction"
256 c 183 g 96 t 3 others
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                170 rgAlaAlaValCysThrArg------183
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           87 IleGlnMetTyrThrAsnValAspLysAspLeuValGlyTrpProAlaProGlnGlySer 106
                                                                                                                                                                                                                                                                                                                                                                                                          117 SerAspLeuTyrLeuValThrArgHisAlaAspValIleProValArgArgGlyAsp 136
                                                                                                                                                                                                                                                                                                                                                                                                                                310 CACGAC-----CCACAAGCCCACCCG------CCCACAGCCACCGCGACTTCC 351
                                                                                                                                                                                                                                                                                                                                                                                                                                                      137 SerArgGlySer-----LeuLeuSerProArgProIleSerTyrLeuLysGlySerSer 154
                                                                                                                                                                                                                                                                                                                               532 CAACCGTGCCCATCGTCAACCTCAAGGACGAACTGCTTTCCA 575
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         184 --PhelleProValGluSerLeuGluThrThrMetArgSerPro 197
                                                                                                                                                                    641
14
14
56
39
                                                                                                                                                                   Length:
Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                              Gaps:
                                                                                                                                                                                                                                                   US-09-965-594-14 (1-197) x AU127824 (1-641)
  Location/Qualifiers
                                                                                                                                                                    34.3
94.50
39.10%
30.13%
9.16%
                                                                                                                                                                                      Percent Similarity:
Best Local Similarity:
Query Match:
                                                                                                                            103
                                                                                                                                                           Alignment Scores:
                                                                                                                          BASE COUNT
ORIGIN
                                                                                                                                                                        Pred. No.:
FEATURES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              9
6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                             a
                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                              QQ
                                                                                                                                                                                                                                                                                                                                                                                                     õ
                                                                                                                                                                                                                                                                                                                                                                                                                              a
                                                                                                                                                                                                                                                                                                                                                                                                                                                      õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                        8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                        ò
```

Search completed: August 31, 2003, 04:27:28 Job time : 1914.31 secs

```
Hepatitis C virus
                                                                                                                                                                                          Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1: /SIDSI/gogdata/geneseq/geneseqp-embl/AA1980.DAT:*
2: /SIDSI/gogdata/geneseq/geneseqp-embl/AA1981.DAT:*
3: /SIDSI/gogdata/geneseqy-embl/AA1981.DAT:*
4: /SIDSI/gogdata/geneseqy-embl/AA1981.DAT:*
5: /SIDSI/gogdata/geneseqy-embl/AA1981.DAT:*
6: /SIDSI/gogdata/geneseqy-embl/AA1986.DAT:*
7: /SIDSI/gogdata/geneseqy-embl/AA1986.DAT:*
8: /SIDSI/gogdata/geneseqy-embl/AA1986.DAT:*
8: /SIDSI/gogdata/geneseqy-embl/AA1980.DAT:*
8: /SIDSI/gogdata/geneseqy-embl/AA1980.DAT:*
9: /SIDSI/gogdata/geneseqy-embl/AA1980.DAT:*
10: /SIDSI/gogdata/geneseqy-embl/AA1980.DAT:*
11: /SIDSI/gogdata/geneseqy-embl/AA1990.DAT:*
12: /SIDSI/gogdata/geneseqy-embl/AA1990.DAT:*
13: /SIDSI/gogdata/geneseqy-embl/AA1990.DAT:*
14: /SIDSI/gogdata/geneseqy-embl/AA1990.DAT:*
15: /SIDSI/gogdata/geneseqy-embl/AA1990.DAT:*
16: /SIDSI/gogdata/geneseqy-embl/AA1990.DAT:*
17: /SIDSI/gogdata/geneseqy-embl/AA1990.DAT:*
18: /SIDSI/gogdata/geneseqy-embl/AA1990.DAT:*
18: /SIDSI/gogdata/geneseqy-embl/AA1990.DAT:*
19: /SIDSI/gogdata/geneseqy-embl/AA1990.DAT:*
20: /SIDSI/gogdata/geneseqy-embl/AA1990.DAT:*
21: /SIDSI/gogdata/geneseqy-embl/AA1990.DAT:*
22: /SIDSI/gogdata/geneseqy-embl/AA1990.DAT:*
23: /SIDSI/gogdata/geneseqy-embl/AA1990.DAT:*
24: /SIDSI/gogdata/geneseqy-embl/AA1990.DAT:*
25: /SIDSI/gogdata/geneseqy-embl/AA1990.DAT:*
26: /SIDSI/gogdata/geneseqy-embl/AA1990.DAT:*
27: /SIDSI/gogdata/geneseqy-embl/AA1990.DAT:*
28: /SIDSI/gogdata/geneseqy-embl/AA1990.DAT:*
28: /SIDSI/gogdata/geneseqy-embl/AA2000.DAT:*
28: /SIDSI/gogdata/geneseqy-embl/AA2000.DAT:*
28: /SIDSI/gogdata/geneseqy-embl/AA2000.DAT:*
29: /SIDSI/gogdata/geneseqy-embl/AA2000.DAT:*
21: /SIDSI/gogdata/geneseqy-embl/AA2000.DAT:*
22: /SIDSI/gogdata/geneseqy-embl/AA2000.DAT:*
23: /SIDSI/gogdata/geneseqy-embl/AA2000.DAT:*
                                                                                                                                                                                                                                                                                                                               US-09-965-594-16
1032
1 MKKKGSVVIVGRINLSGDTA......VAKAVDFIPVESLETTMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Description
GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Total number of hits satisfying chosen parameters:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1107863 segs, 158726573 residues
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SUMMARIES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               AAB15222
AAB15221
AAB15223
AAB15224
AAB15226
AAB15226
AAB15226
AAB15226
AAB15226
AAB15226
                                                                                                                                     OM protein - protein search, using sw model
                                                                                                                                                                                                 August 30, 2003, 17:42:58
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Gapop 10.0 , Gapext 0.5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Minimum DB seq length: 0
Maximum DB seq length: 2000000000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   DB
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    %
Query
Match Length D
                                                                                                                                                                                                                                                                                                                                                                                                                                                                        BLOSUM62
                                                                                                                                                                                                                                                                                                                                                                      Perfect score:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Scoring table:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Score
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ٠.
                                                                                                                                                                                                                                                                                                                                                                                                      Sequence:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Searched:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Database
                                                                                                                                                                                                        Run on:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Result
```

HCV NS4A-NS3 COMP1	NS4A-NS3 NS4A-NS3 NS4A-NS3 NS4A-NS3 NS4A-NS3 NS4A-NS3 NS4A-NS3 NS4A-NS3	HCV NS4A-NS3 COMp1	NS46 NS3 NS3 prote Atitis C v Atitis C v Genomic a V150-encod	Protein sequence o Peptide encoded by Sequence encoded i Sequence encoded i Peptide encoded by Protein encoded by
AAY24947 AAY24941 AAY24942 AAY17880 AAY24945 AAY24946	AAY24946 AAY17844 AAY17884 AAY17879 AAY17878 AAY17878 AAY17890	AAY17882 AAY17883 AAY17877 AAY17877 AAY17881 AAY17881	AAN 17 90 AAN 17 99 ABY 11 460 ABY 10 20 AAR 90 20 AAR 92 23 AAR 92 21	AAP90158 AAP90164 AAP92047 AAP92050 AAP90208
200000	00000000	22222222	000000000000000000000000000000000000000	10 10 10 710 710
665 665 665 665 665	665 216 216 216 665 671	216 216 216 215 215 216	631 3011 3011 3011 3011 687 1648	1786 2261 2301 2436 2772
			866.11 866.11 865.11 885.77	
908.5 907.5 904.5 904.5	9000 9001.5 9000.5 900.5 900.5	897.5 897.5 896.5 893.5 893.5 893.5	99 89 89 89 89 89 89 89 89 89 89 89 89 8	88888888888888888888888888888888888888
0 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1	110 110 110 110 110 110 110 110 110 110	2000000 200000000000000000000000000000	1 W W W W W W W W W W W W W W W W W W W	44444 04444

Ē	
-	
ח	
J	

```
Modified hepatitis C virus (HCV) NS3 protease comprising at least 1
                                                                                                     Hepatitis; NS3 protease; viral replication; chronic liver disease; liver failure; liver cancer; mutant; mutein.
                                                                                                                                                                                                                                                                           Goldfarb V;
                                                                                 Hepatitis C virus NS4A-NS3 fusion protease #4.
                                                                                                                                                                                                                                                                          Wittekind M, Weinheimer S, Zhang Y,
                     AAB15222 standard; protein; 197 AA.
                                                                                                                                                                                                                                                      (BRIM ) BRISTOL-MYERS SQUIBB CO.
                                                                                                                                                                                                           06-JAN-2000; 2000WO-US00345.
                                                                                                                                                                                                                                990S-0115271.
                                                             19-DEC-2000 (first entry)
                                                                                                                                                                                                                                                                                              WPI; 2000-465976/40.
N-PSDB; AAA73331.
                                                                                                                                    Hepatitis C virus
Synthetic.
                                                                                                                                                                  WO200040707-A1.
                                                                                                                                                                                                                                08-JAN-1999;
                                                                                                                                                                                       13-JUL-2000
                                        AAB15222
RESULT 1
```

```
ö
                                                                                    The present sequence is a mutated version of a fusion protein created using the Hepatitis C virus (HCV) NS3 and NS4A protease enzymes. These proteins are both essential for the replication of the virus, acting to cleave its replicative proteins from the polyprotein produced from the HCV genome. Inhibitors of the two proteins should be effective as antiviral treatments of HCV infection. This is useful as HCV can lead to chronic liver disease such as cirrhosis, liver failure and liver cancer. The present invention concerns a number of NS3 mutants and NS3-NS4A fusion proteins which can be used to identify inhibitors of this type, as
                                                                                                                                                                                                                                                                                                                                                                                                                                              TCINGVCWTVYHGAGTRTIASPKGPVTQMYTNVDKDLVGWQAPQGSRSLTPCTCGSSDLY 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       LVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVCTRGVAK 180
                                                                                                                                                                                                                                                                                                                                                                                               9
                                                                                                                                                                                                                      well as enabling structural studies of the protease and protease:inhibitor complexes. This sequence contains the alpha-helix0-1
substitution of a hydrophobic alpha-helix 0 amino acid to a hydrophilic amino acid, useful for screening inhibitors that may treat hepatitis C
                                                                                                                                                                                                                                                                                                                                                                                      1 MKKKGSVVIVGRINLSGDTAYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVSTATQTFLA
                                                                                                                                                                                                                                                                                                                                         Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Hepatitis; NS3 protease; viral replication; chronic liver disease;
liver failure; liver cancer; mutant; mutein.
                                                                                                                                                                                                                                                                                                                                         ;
0
                                                                                                                                                                                                                                                                                                            DB 21; Length 197;
                                                                                                                                                                                                                                                                                                                                        Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Goldfarb V;
                                                                                                                                                                                                                                                                                                100.0%; Score 1032; DB 21; 100.0%; Pred. No. 6e-101; Wiematches 0;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Hepatitis C virus NS4A-NS3 fusion protease #3.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Wittekind M, Weinheimer S, Zhang Y,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    AAB15221 standard; protein; 197 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                8
                                                            Claim 23; Fig 14; 66pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                181 AVDFIPVESLETTMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                AVDFIPVESLETTMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                (BRIM ) BRISTOL-MYERS SQUIBB
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       2000WO-US00345
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   99US-0115271
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (first entry)
                                                                                                                                                                                                                                                                                                                                         Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        WPI; 2000-465976/40.
                                                                                                                                                                                                                                                                                                                           Similarity
                                                                                                                                                                                                                                                                               197 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       N-PSDB; AAA73330
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          40200040707-A1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    06-JAN-2000;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   08-JAN-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             19-DEC-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         13-JUL-2000
                                                                                                                                                                                                                                                                                                                                      197;
                                                                                                                                                                                                                                                                                  Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Synthetic
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                AAB15221;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          181
                                                                                                                                                                                                                                                                                                                                                                                                                              61
                                                                                                                                                                                                                                                                                                            Query Match
                                                                                                                                                                                                                                                                                                                             Local
                                                                                                                                                                                                                                                                                                                           Best Loca
Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       RESULT 2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   a
                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                              ç
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
```

```
ô
                                                                                                                The present sequence is a mutated version of a fusion protein created using the Hepatitis C virus (HCV) NG3 and NS4A protease enzymes. These proteins are both essential for the replication of the virus, acting to cleave its replicative proteins from the polyprotein produced from the HCV genome. Inhibitors of the two proteins should be effective as antiviral treatments of HCV infection. This is useful as HCV can lead to chronic liver disease such as cirrhosis, liver failure and liver cancer. The present invention concerns a number of NS3 mutants and NS3-NS4A fusion proteins which can be used to identify inhibitors of this type, as well as enabling structural studies of the protease and protease:inhibitor complexes. This sequence contains the alpha-helixo-1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    61 TCINGVCMTVYHGAGTRIIASPKGPVIQMYINVDKDLVGMPAPQGSRSLTPCTCGSSDLY 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              61 TCINGVCWTVYHGAGTRTIASPKGPVTQMYTNVDKDLVGWQAPQGSRSLTPCTCGSSDLY 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                121 LVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVCTRGVAK 180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                Modified hepatitis C virus (HCV) NS3 protease comprising at least 1 substitution of a hydrophobic alpha-helix 0 amino acid to a hydrophilic amino acid, useful for screening inhibitors that may treat hepatitis C
                                                                                                                                                                                                                                                                                                                                                                                                                                              1 MKKKGSVVIVGRINLSGDTAYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVSTATQTFLA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       protease; viral replication; chronic liver disease; liver cancer; mutant; mutein.
                                                                                                                                                                                                                                                                                                                                                                                                            ö
                                                                                                                                                                                                                                                                                                                                                                           Score 1015; DB 21; Length 197; Pred. No. 3.8e-99;
                                                                                                                                                                                                                                                                                                                                                                                                              Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Goldfarb V;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Hepatitis C virus NS4A-NS3 fusion protease #5.
                                                                                                                                                                                                                                                                                                                                                                                                           0; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Wittekind M, Weinheimer S, Zhang Y,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Ź
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      AAB15223 standard; protein; 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  8
                                                                                   Claim 23; Fig 13; 66pp; English
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  181 AVDFIPVESLETTMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  181 AVDFIPVESLETTMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  (BRIM ) BRISTOL-MYERS SQUIBB
                                                                                                                                                                                                                                                                                                                                                                         98.48;
98.58;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 99US-0115271
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               06-JAN-2000; 2000WO-US00345
                                                                                                                                                                                                                                                                                                                                                                                         Best Local Similarity 98.5
Matches 194; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    WPI; 2000-465976/40.
                                                                                                                                                                                                                                                                                                                                          197 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      N-PSDB; AAA73332
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           liver failure;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Hepatitis; NS3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           WO200040707-A1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 08-JAN-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      19-DEC-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             13-JUL-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Synthetic.
                                                                                                                                                                                                                                                                                                                                            Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      AAB15223;
                                                                                                                                                                                                                                                                                                                                                                           Query Match
Best Local S
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      AAB15223
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
```

```
The present sequence is a mutated version of a fusion protein created using the Hepatitis C virus (HCV) NG3 and NS4A protease enzymes. These proteins are both essential for the replication of the virus, acting to cleave its replicative proteins from the polyprotein produced from the HCV genome. Inhibitors of the two proteins should be effective as antiviral treatments of HCV infection. This is useful as HCV can lead to chronic liver disease such as cirrhosis, liver failure and liver cancer. The present invention concerns a number of NS3 mutants and NS3-NS4A (Fusion proteins which can be used to identify inhibitors of this type, as well as enabling structural studies of the protease and protease:Inhibitor complexes. This sequence contains the alpha-helix0-1
                                                                                                                                                                                                                                                                                                                                                                              0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                     61 TCINGVCWTVYHGAGTRTIASPRGPVTQMYTNVDKDLVGWQAPQGSRSLTPCTCGSSDLY 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   LVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVCTRGVAK 180
                                                                                                                                                                                                                                                                                                                                                                                                                                       9
                                                                                                                                                                                                                                                                                                                                                                                                           9
            Modified hepatitis C virus (HCV) NS3 protease comprising at least 1 substitution of a hydrophobic alpha-helix 0 amino acid to a hydrophilic amino acid, useful for screening inhibitors that may treat hepatitis C
                                                                                                                                                                                                                                                                                                                                                                                                                          1 MKKKGSVVIVGRINLSGDTAYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVSTATQTFLA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      61 TSINGVLWTVYHGAGTRTIASPKGPVTQMYTNVDKDLVGWQAPQGSRSLTPCTGGSSDLY
                                                                                                                                                                                                                                                                                                                                                                                Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Hepatitis; NS3 protease; viral replication; chronic liver disease;
liver failure; liver cancer; mutant; mutein.
                                                                                                                                                                                                                                                                                                                                                                              ö
                                                                                                                                                                                                                                                                                                                                               Length 197;
                                                                                                                                                                                                                                                                                                                                                                             Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Goldfarb V;
                                                                                                                                                                                                                                                                                                                                               Score 1002; DB 21;
Pred. No. 9e-98;
0; Mismatches 3;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Hepatitis C virus NS4A-NS3 fusion protease #6.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Zhang Y,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Ş
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    AAB15224 standard; protein; 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           (BRIM ) BRISTOL-MYERS SQUIBB CO.
                                                                                        23; Fig 15; 66pp; English
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            181 AVDFIPVESLETTMRSP 197
                                                                                                                                                                                                                                                                                                                                                97.1%;
98.5%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Wittekind M, Weinheimer S,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             06-JAN-2000; 2000WO-US00345
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           AVDFIPVESLETIMRSP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              (first entry)
                                                                                                                                                                                                                                                                                                                                                              Best Local Similarity 98.5
Matches 194; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     WPI; 2000-465976/40
                                                                                                                                                                                                                                                                                                                  197 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 WO200040707-A1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             08-JAN-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              19-DEC-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               13-JUL-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Synthetic.
                                                                                                                                                                                                                                                                                                                     Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                AAB15224;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          181
                                                                                                                                                                                                                                                                                                                                                Query Match
Best Local S
                                                                                                                                                                                                                                                                                          variant
                                                                                          Claim
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    RESULT 4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ç
                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                       õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        a
```

```
The present sequence is a mutated version of a fusion protein created using the Hepatitis C virus (HCV) NG3 and NG4A protease enzymes. These proteins are both essential for the replication of the virus, acting to cleave its replicative proteins from the polyprotein produced from the HCV genome. Inhibitors of the two proteins should be effective as antiviral treatments of HCV infection. This is useful as HCV can lead to chronic liver disease such as cirthosis, liver failure and liver cancer. The present invention concerns a number of NG3 mutants and NG3-NG4A well as enabling structural studies of the protease and protease: Inhibitor complexes. This sequence contains the alpha-helix0-7
                                                                                                                                                                                                                                                                                                                                                              0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  120
                                                                                                                                                                                                                                                                                                                                                                                                                                                             9
                                                                                                                                                                                                                                                                                                                                                                                                                  9
                        Modified hepatitis C virus (HCV) NS3 protease comprising at least 1 substitution of a hydrophobic alpha-helix 0 amino acid to a hydrophilic amino acid, useful for screening inhibitors that may treat hepatitis C
                                                                                                                                                                                                                                                                                                                                                                                       1 MKKKGSVVIVGRINLSGDTAYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVSTATQTFLA
                                                                                                                                                                                                                                                                                                                                                                                                       TCINGVCWTVYHGAGTRTIASPRGPVTQMYTNVDKDLVGWQAPQGSRSLTPCTCGSSDLY
                                                                                                                                                                                                                                                                                                                                                                Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Hepatitis; NS3 protease; viral replication; chronic liver disease; liver failure; liver cancer; mutant; mutein.
                                                                                                                                                                                                                                                                                                                                                              ö
                                                                                                                                                                                                                                                                                                                                  Length 197;
                                                                                                                                                                                                                                                                                                                                                              Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Goldfarb V;
                                                                                                                                                                                                                                                                                                                                 95.9%; Score 990; DB 21; 97.0%; Pred. No. 1.7e-96; tive 2; Mismatches 4;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Hepatitis C virus NS4A-NS3 fusion protease #2.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Zhang Y,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Ź
                                                                                             Claim 23; Fig 16; 66pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           AAB15220 standard; protein; 195
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          BRIM ) BRISTOL-MYERS SQUIBB CO
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          AVDFIPVESLETIMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Weinheimer S,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 06-JAN-2000; 2000WO-US00345
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             990S-0115271
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                19-DEC-2000 (first entry)
                                                                                                                                                                                                                                                                                                                                                Best Local Similarity 97.0
Matches 191; Conservative
                                                                                                                                                                                                                                                                                                        197 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Hepatitis C virus
N-PSDB; AAA73333.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          WO200040707-A1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Wittekind M,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           08-JAN-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     13-JUL-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Synthetic
                                                                                                                                                                                                                                                                                                                                                                                                                                                 61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       AAB15220
                                                                                                                                                                                                                                                                                                          Sequence
                                                                                                                                                                                                                                                                                                                                  Query Match
                                                                                                                                                                                                                                                                                variant
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                RESULT 5
ð
                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                              ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                       a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                d
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  a
```

us-09-965-594-16.rag

```
AAB15226
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          õ
                                                                                                                                                                                                                                                                                                                                                                      ద
                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                              ద
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            셤
                                                                                                             The present sequence is a mutated version of a fusion protein created using the Hepatitis C virus (HeV) NS3 and NS4A protease enzymes. These proteins are both essential for the replication of the virus, acting to cleave its replicative proteins from the polyprotein produced from the HVV genome. Inhibitors of the two proteins should be effective as antiviral treatments of the two proteins should be effective as antiviral treatments of HCV infection. This is useful as HCV can lead to chronic liver disease such as cirrhosis, liver failure and liver cancer. The present invention concerns a number of NS3 mutants and NS3-NS4A fusion proteins which can be used to identify inhibitors of this type, as well as enabling structural studies of the protease and protease: Inhibitor complexes. This sequence contains the alpha-helix0-1
                                                                                                                                                                                                                                                                                                                                                                                             TCINGVCWTVYHGAGTRTIASPKGPVTQMYTNVDKDLVGWQAPQGSRSLTPCTCGSSDLY 120
                                                                                                                                                                                                                                                                                                                                                                                                                              LVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVCTRGVAK 180
                                                                                                                                                                                                                                                                                                                                               1 HKKKGSVVIVGRIVLNG--AYAQQTRGEGCCCTTSQTGRDKNQVEGEVQIVSTAAQTFLA 58
                                 Modified hepatitis C virus (HCV) NS3 protease comprising at least 1 substitution of a hydrophobic alpha-helix 0 amino acid to a hydrophilic amino acid, useful for screening inhibitors that may treat hepatitis {\sf C}
                                                                                                                                                                                                                                                                                                                                    1 MRKKGSVVIVGRINLSGDTAYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVSTATQTFLA
                                                                                                                                                                                                                                                                                                               Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Hepatitis; NS3 protease; viral replication; chronic liver disease; liver failure; liver cancer; mutant; mutein.
                                                                                                                                                                                                                                                                                                               ;
                                                                                                                                                                                                                                                                                        Length 195;
                                                                                                                                                                                                                                                                                                              4; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Goldfarb V;
                                                                                                                                                                                                                                                                                       Score 981; DB 21;
Pred. No. 1.5e-95;
1; Mismatches 4;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Hepatitis C virus NS4A-NS3 fusion protease #7.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Zhang Y,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       AAB15225 standard; protein; 197 AA.
                                                                                          Claim 23; Fig 12; 66pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (BRIM ) BRISTOL-MYERS SQUIBB CO.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       AVDFIPVESLETTMRSP 195
                                                                                                                                                                                                                                                                                                                                                                                                                                                                          AVDFIPVESLETTMRSP 197
                                                                                                                                                                                                                                                                                       95.1%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Weinheimer S,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       99US-0115271.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 06-JAN-2000; 2000WO-US00345
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (first entry)
                                                                                                                                                                                                                                                                                                              190; Conservative
WPI; 2000-465976/40
                                                                                                                                                                                                                                                                                       Query Match
Best Local Similarity
                                                                                                                                                                                                                                                                 195 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Hepatitis C virus.
Synthetic.
         N-PSDB; AAA73329
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  WO200040707-A1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Wittekind M,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       08-JAN-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     19-DEC-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         13-JUL-2000
                                                                                                                                                                                                                                                                                                                                                                                 61
                                                                                                                                                                                                                                                                                                                                                                                                                            121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                         181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               179
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              AAB15225;
                                                                                                                                                                                                                                                                  Seguence
                                                                                                                                                                                                                                             variant
                                                                                                                                                                                                                                                                                                              Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               AAB15225
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   RESULT
 õ
                                                                                                                                                                                                                                                                                                                                                         셤
                                                                                                                                                                                                                                                                                                                                                                                õ
                                                                                                                                                                                                                                                                                                                                                                                                    q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   å
                                                                                                                                                                                                                                                                                                                                                                                                                                                  8
```

```
The present sequence is a mutated version of a fusion protein created using the Hepatitis C virus (HVV) NG3 and NS4A protease enzymes. These proteins are both essential for the replication of the virus, acting to cleave its replicative proteins from the polyprotein produced from the HCV genome. Inhibitors of the two proteins should be effective as antiviral treatments of the two proteins should be effective as antiviral treatments of HCV infection. This is useful as HCV can lead to chronic liver disease such as cirthosis, liver failure and liver cancer. The present invention concerns a number of NS3 mutants and NS3-NS4A fusion proteins which can be used to identify inhibitors of this type, as well as enabling structural studies of the protease and protease: inhibitor complexes. This sequence contains the alpha-helix0-7
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 LVTRHADVIPVRRKGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVCTRGVAK 180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              9
                                                                              Modified hepatitis C virus (HCV) NS3 protease comprising at least 1 substitution of a hydrophobic alpha-helix 0 amino acid to a hydrophilic amino acid, useful for screening inhibitors that may treat hepatitis C
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              TCINGVCWTVYHGAGTRIIASPKGPVTQMYINVDKDLVGWQAPQGSRSLTPCTCGSSDLY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1 MKKKGSVVIVGRINLSGDTAYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVSTATQTFLA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      protease; viral replication; chronic liver disease; liver cancer; mutant; mutein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Score 980; DB 21; Length 197;
Pred. No. 1.9e-95;
2; Mismatches 5; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Hepatitis C virus NS4A-NS3 fusion protease #8.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      AAB15226 standard; protein; 197 AA.
                                                                                                                                                                                                                                   Claim 23; Fig 17; 66pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (BRIM ) BRISTOL-MYERS SQUIBB CO.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                181 AVDFIPVESLETIMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  95.0%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          99US-0115271.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                06-JAN-2000; 2000WO-US00345
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Query Match
Best Local Similarity 96.4
Matches 190; Conservative
2000-465976/40.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      197 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Hepatitis C virus
Synthetic.
                           N-PSDB; AAA73334
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Hepatitis; NS3
liver failure;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             WO200040707-A1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          08-JAN-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      19-DEC-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          13-JUL-2000.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            AAB15226;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  variant.
```

us-09-965-594-16.rag

```
Wittekind M,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               W09928482-A2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          24-NOV-1998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                28-JUL-1998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         28-NOV-1997;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   10-JUN-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    AAY24943;
                                                                                                                                                                                                                                                                         Seguence
                                                                                                                                                                                                                                                                                              Query Match
                                                                                                                                                                                                                                                                                                         Local
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   AAY24943
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          RESULT
   õ
                                                                                                                                                                                                                                                                                                                                                           셤
                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                    셤
                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         The present sequence is a mutated version of a fusion protein created using the Hepatitis C virus (HCV) NS3 and NS4A protease enzymes. These proteins are both essential for the replication of the virus, acting to cleave its replicative proteins from the polyprotein produced from the HCV genome. Inhibitors of the two proteins should be effective as antiviral treatments of HCV infection. This is useful as HCV can lead to chronic liver disease such as cirrhosis, liver failure and liver cancer. The present invention concerns a number of NS3 mutants and NS3-NS4A fusion proteins which can be used to identify inhibitors of this type, as well as enabling structural studies of the protease and protease:inhibitor complexes. This sequence contains the alpha-helix0 wild-type sequence.
                                                                                                                                                                                                                                                                                                                  ö
                                                                                                                                                                                                                                                                                                                                                                                 61 TCINGVCWTVYHGAGTRIIASPKGPVTQMYINVDKDLVGWQAPQGSRSLTPCTCGSSDLY 120
                                                                                                                                                                                                                                                                                                                                                                                                      120
                                                                                                                                                                                                                                                                                                                                       9
                                                                                                                                                                                                                                                                                                                                                            9
                                                     Modified hepatitis C virus (HCV) NS3 protease comprising at least 1 substitution of a hydrophobic alpha-helix 0 amino acid to a hydrophilic amino acid, useful for screening inhibitors that may treat hepatitis C
                                                                                                                                                                                                                                                                                                                                                   1 MKKKGSVVIVGRINLSGDFAYAQQFRGEEGCQETSQFGRDKNQVEGEVQIVSTATQFFLA
                                                                                                                                                                                                                                                                                                                                                                                                                           LVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVCTRGVAK
                                                                                                                                                                                                                                                                                                                  Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Hepatitis; NS3 protease; viral replication; chronic liver disease; liver failure; liver cancer.
                                                                                                                                                                                                                                                                                                                  ;
0
                                                                                                                                                                                                                                                                                          94.6%; Score 976; DB 21; Length 197; 95.9%; Pred. No. 5.1e-95; 1.ve 0; Mismatches 8; Indels
 Goldfarb V;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Hepatitis C virus NS4A-NS3 fusion protease #1.
Weinheimer S, Zhang Y,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              $
                                                                                                           Example 5; Fig 18; 66pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (BRIM ) BRISTOL-MYERS SQUIBB CO.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             AAB15212 standard; protein; 195
                                                                                                                                                                                                                                                                                                                                                                                                                                                                               181 AVDFIPVESLETTMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    99US-0115271
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             16-JAN-2000; 2000WO-US00345
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        19-DEC-2000 (first entry)
                                                                                                                                                                                                                                                                                                       Best_Local Similarity 95.9
Matches 189; Conservative
                     WPI; 2000-465976/40.
N-PSDB; AAA73335.
                                                                                                                                                                                                                                                                       197 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 WO200040707-A1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    08-JAN-1999;
Wittekind M,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         13-JUL-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Synthetic.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   AAB15212;
                                                                                                                                                                                                                                                                        Sequence
                                                                                                                                                                                                                                                                                            Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        RESULT
AAB1521
셤
                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                           ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        a
```

```
61 TCINGVCWTVYHGAGTRIIASPKGPVTQMYTNVDKDLVGWQAPQGSRSLTPCTCGSSDLY 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               8
                                                                                                                                                                                                                                                                                                                                                                     The present sequence is a fusion protein created using the Hepatitis C virus (HCV) N33 and N84A protease enzymes. These proteins are both essential for the replication of the virus, acting to cleave its replicative proteins from the polyprotein produced from the HCV genome. Inhibitors of the two proteins should be effective as antiviral treatments of HCV infection. This is useful as HCV can lead to chronic liver disease such as cirrhosis, liver failure and liver cancer. The present invention concerns a number of MS3 mutents and MS3-MS4 fusion proteins which can be used to identify inhibitors of this type, as well as enabling structural studies of the protease and protease:Inhibitor
                                                                                                                                                      Modified hepatitis C virus (HCV) NS3 protease comprising at least 1 substitution of a hydrophobic alpha-helix 0 amino acid to a hydrophilic amino acid, useful for screening inhibitors that may treat hepatitis C
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1 MKKKGSVVIVGRINLSGDTAYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVSTATQTFLA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                HCV; hepatitis C virus; single chain recombinant complex; linker; NS3 protease; NS4A cofactor; NS4A-NS3 complex; serine protease; hydrophobic domain; covalent complex; detection; inhibitor.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    5.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   91.3%; Score 942; DB 21; Length 195; 93.9%; Pred. No. 2e-91;
Goldfarb
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1; Mismatches
Zhang Y,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Ź
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        HCV NS4A-NS3 complex SEQ ID NO:14.
                                                                                                                                                                                                                                                                                                                 Example 2; Fig 10; 66pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                AAY24943 standard; Protein; 665
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            AVDFIPVESLETTMRSP 195
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             98US-0094331.
Weinheimer S,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  98WO-US24528
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          07-SEP-1999 (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Matches 185; Conservative
                                                           WPI; 2000-465976/40.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        195 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Hepatitis C virus
Synthetic.
                                                                                           N-PSDB; AAA73328
```

```
(SCHE ) SCHERING CORP.
   (SCHE ) SCHERING CORP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     W09928482-A2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  24-NOV-1998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                28-NOV-1997;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          07-SEP-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       28-JUL-1998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           10-JUN-1999
                         Malcolm BA,
                                                                                                                                                                                                                                                                Seguence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               à
                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                            à
                                                                                                                                                                                                                                                                                                                                                                                              q
                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                            8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                   à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        a
                                                                                                                                                                                                                                                                                                                                                                                                                                           62 CINGVCWTVYHGAGTRTIASPKGPVTQMYTNVDKDLVGWQAPQGSRSLTPCTCGSSDLYL 121
                                                                                                                                                                                                                                                                                                                                                                                                                               122 VTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVCTRGVAKA 181
                                                                                                                       The present invention describes a covalent hepatitis C virus (HCV) NS4A-NS3 complex comprising a central hydrophobic domain of native HCV NS4A-NS3 complex comprising a central hydrophobic domain of native HCV NS3 berine protease domain, where the hydrophobic domain of native HCV NS4A peptide is tethered by the linker to the amino terminus of the HCV NS3 protease domain. The present complex. The covalent NS4A-NS3 complexes are useful for structural determination and determination of mode of binding of HCV inhibitors by NMR spectroscopy. They can also be used for detecting inhibitors of the protease activity, the helicase activity and the ATPase activity of NS3. The covalent NS4A-NS3 complexes are more soluble, stable and active than the non-covalent protease-peptide complexes previously available.
                                                                                                                                                                                                                                                                                                                                                               81
                                                                                                                                                                                                                                                                                                                                           61
                                                                                                                                                                                                                                                                                                                                                      22 GSVVIVGRIILSGSGSITAYSQQTRGLLGCKKTSLTGRDKNQVEGEVQVVSTATQSFLAT
                                                                                                                                                                                                                                                                                                                                                                                                GSVVIVGRINLSGD -- - TAYAQQTRGEEGCOETSQTGRDKNQVEGEVQIVSTATQTFLAT
                                                                                                                                                                                                                                                                                                                     3; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              HCV; hepatitis C virus; single chain recombinant complex; linker; NS3 protease; NS4A cofactor; NS4A-NS3 complex; serine protease; hydrophobic domain; covalent complex; detection; inhibitor.
                                                                                                                                                                                                                                                                                            88.3%; Score 911.5; DB 20; Length 665; 86.7%; Pred. No. 1.9e-87; tive 16; Mismatches 7; Indels 3;
                                                                             New hepatitis C virus covalent complexes
                                                                                                     Claim 6; Page 90-92; 211pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         AAY24947 standard; Protein; 665 AA.
                                   Weber PC,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          HCV NS4A-NS3 complex SEQ ID NO:18.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        11:11:11:11:11
202 VDFVPVESMETTMRSP 217
                                                                                                                                                                                                                                                                                                                                                                                                                                                                           182 VDFIPVESLETTMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           98WO-US24528
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                98US-0094331
97US-0067315
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (first entry)
                                                                                                                                                                                                                                                                                                        al Similarity 86.79
170; Conservative
                                   Taremi SS,
           (SCHE ) SCHERING CORP
                                                        WPI; 1999-385385/32
                                                                                                                                                                                                                                                                         665 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              W09928482-A2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   07-SEP-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          24-NOV-1998;
                                   Malcolm BA,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               28-JUL-1998
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          28-NOV-1997
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     10-JUN-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Synthetic.
                                                                                                                                                                                                                                                                                             Query Match
Best Local Si
Matches 1700
                                                                                                                                                                                                                                                                         Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              AAY24947;
                                                                                                                                                                                                                                                                                                                                          S
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    RESULT 10
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               AAX 24947
g
                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          a
                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
```

```
The present invention describes a covalent hepatitis C virus (HCV) NS4A-NS3 complex comprising a central hydrophobic domain of native HCV NS3 earlie protease domain. Where the hydrophobic domain of native HCV NS3 earlie protease domain. Where the hydrophobic domain of native HCV NS4A peptide is tethered by the linker to the amino terminus of the HCV NS3 protease domain. The present sequence represents a specifically claimed example of the above complex. The covalent NS4A-NS3 complexes are useful for structural determination and determination of mode of binding of HCV inhibitors by NMR spectroacopy. They can also be used for detecting inhibitors by protease activity, the helicase activity and the ATPase activity of NS3. The covalent NS4A-NS3 complexes are more soluble, stable and active than the non-covalent protease-peptide complexes previously available.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               62 CINGVCWTVYHGAGTRTIASPKGPVTQMYTNVDKDLVGWQAPQGSRSLTPCTCGSSDLYL 121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        122 VTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVCTRGVAKA 181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         81
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           22 GSVVIVGRIILSGSGSITAYSQQTRGLLGCKKTSLTGRDKNQVEGEVQVVSTATQSFLAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      5 GSVVIVGRINLSGD---TAYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVSTATQTFLAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   HCV; hepatitis C virus; single chain recombinant complex; linker; NS3 protease; NS4A cofactor; NS4A-NS3 complex; serine protease; hydrophobic domain; covalent complex; detection; inhibitor.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          3;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          DB 20; Length 665;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       88.0%; Score 908.5; DB 20; Length
86.2%; Pred. No. 3.9e-87;
Live 17; Mismatches 7; Indels
   ż
                                                                                                                        New hepatitis C virus covalent complexes
                                                                                                                                                                                        Claim 6; Page 100-102; 211pp; English
Weber PC,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 AAY24941 standard; Protein; 665 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          HCV NS4A-NS3 complex SEQ ID NO:12.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            182 VDFIPVESLETIMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             97US-0067315
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      98US-0094331
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      98WO-US24528
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Query Match
Best Local Similarity 86.24
Matches 169; Conservative
Taremi SS,
                                                         WPI; 1999-385385/32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          665 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Hepatitis C virus
Synthetic.
```

1;

```
Malcolm BA, Taremi SS, Weber PC, Yao N;
 Taremi SS,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      W09928482-A2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     24-NOV-1998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            28-JUL-1998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         28-NOV-1997;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        07-SEP-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             10-JUN-1999
 Malcolm BA,
                                                                                                                                                                                                                                               Sednence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                AAY17880;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                AAY17880
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                q
                                                                                                                                                                                                                                                                                                                                                                ó
                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                          à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                               a
                                                                                                                                                                                                                                                                                                      Ϊ;
                                                                                                                                                                                                                                                                                                                                                                         62 CINGVCWTVYHGAGTRTIASPKGPVTQMYTNVDKDLVGWQAPQGSRSLTPCTCGSSDLYL 121
                                                                                                                                                                                                                                                                                                                                                                                      122 VTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVCTRGVAKA 181
                                                                                                                                                                                                                                                                                                                                                                                                                                    The present invention describes a covalent hepatitis C virus (HCV) NS4A-NS3 complex comprising a central hydrophobic domain of native HCV NS4A-NS3 complex, and an HCV NS3 earlie protease domain, where the hydrophobic domain of native HCV NS4A peptide is tethered by the linker to the amino terminus of the HCV NS4A peptide is tethered by the linker to the amino terminus of the HCV NS3 protease domain. The present sequence represents a specifically claimed example of the above complex. The covalent NS4A-NS3 complexes are useful for structural determination and determination of mode of binding of HCV inhibitors by NMR specificacypy. They can also be used for detecting inhibitors of the protease activity, the helicase activity and the ATPase activity of NS3. The covalent NS4A-NS3 complexes are more soluble, stable and active than the non-covalent protease-peptide complexes previously available.
                                                                                                                                                                                                                                                                                                                            5 GSVVIVGRINLSGD---TAYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVSTATQTFLAT 61
                                                                                                                                                                                                                                                                                                                                           22 GSVVIVGRIILSGSGSIIAYSQQTRGLLGCKIISLTGRDKNQVEGEVQVVSTATQSFLAT 81
                                                                                                                                                                                                                                                                                                       Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               HCV; hepatitis C virus; single chain recombinant complex; linker; NS3 protease; NS4A cofactor; NS4A-NS3 complex; serine protease; hydrophobic domain; covalent complex; detection; inhibitor.
                                                                                                                                                                                                                                                                                                      ж
;
                                                                                                                                                                                                                                                                              87.9%; Score 907.5; DB 20; Length 665; 86.7%; Pred. No. 5e-87;
                                                                                                                                                                                                                                                                                                      8; Indels
                                                                                                                                                                                                                                                                                                     15; Mismatches
            Yao N;
                                                         New hepatitis C virus covalent complexes
                                                                               Claim 6; Page 85-87; 211pp; English
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     AAY24942 standard; Protein; 665 AA
            Weber PC,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         HCV NS4A-NS3 complex SEQ ID NO:13.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 202 VDFVPVESMETTMRSP 217
                                                                                                                                                                                                                                                                                                                                                                                                                                                                     182 VDFIPVESLETTMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      98US-0094331.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              98WO-US24528
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  (first entry)
                                                                                                                                                                                                                                                                                                     Matches 170; Conservative
            Taremi SS,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          (SCHE ) SCHERING CORP
                                                                                                                                                                                                                                                                                           Best Local Similarity
                               WPI; 1999-385385/32
                                                                                                                                                                                                                                                         665 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 WO9928482-A2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                24-NOV-1998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      28-JUL-1998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  07-SEP-1999
            Malcolm, BA,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        10-JUN-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   28-NOV-1997
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Synthetic
                                                                                                                                                                                                                                                          Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           AAY24942;
                                                                                                                                                                                                                                                                               Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                RESULT 12
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           AAX 24942
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                  a
                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                        ð
```

```
62 CINGVCWTVYHGAGTRTIASPKGPVTQMYTNVDKDLVGWQAPQGSRSLTPCTCGSSDLYL 121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         122 VTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVCTRGVAKA 181
                                                                                                                                                                                                                                                                                                                  NS4A-NS3 complex comprising a central hydrophobic domain of native HCV NS4A-NS3 complex comprising a central hydrophobic domain of native HCV NS3 earine protease domain, where the hydrophobic domain of native HCV NS4A peptide is tethered by the linker to the amino terminas of the HCV NS4 protease domain. The present soquence erepresents a specifically claimed example of the above complex. The covalent NS4A-NS3 complexes are useful for structural determination and determination of mode of binding of HCV inhibitors by MR specificacy. They can also be used for detecting inhibitors of the protease activity, the helicase activity and the ATPase activity of NS3 The covalent NS4A-NS3 complexes are more soluble, stable and active than
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     81
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   22 GSVVIVGRIILSGSGSITAYSQQTRGLLGCIKTSLTGRDKNQVEGEVQVVSTATQSFLAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             5 GSVVIVGRINLSGD --- TAYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVSTATQTFLAT
                                                                                                                                                                                                                                                                                       present invention describes a covalent hepatitis C virus (HCV)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   HCV; hepatitis C virus; single chain recombinant complex; linker; NS3 protease; NS4A cofactor; NS4A-NS3 complex; serine protease; hydrophobic domain; covalent complex; detection; inhibitor.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     the non-covalent protease-peptide complexes previously available
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ٠.
ش
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     DB 20; Length 665;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  8;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Query Match

87.9%; Score 907.5; DB
Best Local Similarity 86.7%; Pred. No. 5e-87;
Matches 170; Conservative 15; Mismatches
ż
                                                                                                                                           New hepatitis C virus covalent complexes
                                                                                                                                                                                                            Claim 6; Page 88-90; 211pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           AAY17880 standard; Protein; 216 AA.
Weber PC,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               HCV NS4A-NS3 complex SEQ ID NO:4.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     182 VDFIPVESLETIMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           VDFVPVESMETTMRSP 217
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                98WO-US24528
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    98US-0094331
97US-0067315
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (SCHE ) SCHERING CORP
                                                                    WPI; 1999-385385/32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             665 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Repatitis C virus
```

œ

```
Claim 6; Page 95-97; 211pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Ź
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Weber PC,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      HCV NS4A-NS3 complex SEQ ID NO:17.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            AAY24946 standard; Protein; 665
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             182 VDFIPVESLETIMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 202 VDFVPVESMETTMRSP 217
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        98WO-US24528
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      98US-0094331
97US-0067315
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       (first entry)
                                                                                                                                                                                                                                                                                                                                            Best Local Similarity 86.2
Matches 169; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Taremi SS,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       (SCHE ) SCHERING CORP
WPI; 1999-385385/32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 WPI; 1999-385385/32
                                                                                                                                                                                                                                                                                               665 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           W09928482-A2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      28-JUL-1998;
28-NOV-1997;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       24-NOV-1998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Malcolm BA,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       07-SEP-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         10-JUN-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Synthetic
                                                                                                                                                                                                                                                                                                 Seguence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       AAY24946;
                                                                                                                                                                                                                                                                                                                           Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        AAY24946
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                       õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ద
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           셤
                                                                                                                    NS4A-NS3 complex comprising a central hydrophobic domain of native HCV NS4A peptide, a linker, and an HCV NS3 berine protease domain, where the hydrophobic domain of native HCV NS4A peptide is tethered by the linker to the amino terminus of the HCV NS3 protease domain. The present sequence represents a specifically claimed example of the above complex. The covalent NS4A-NS3 complexes are useful for structural determination and determination of mode of binding of HCV inhibitors by NMR spectroscopy. They can also be used for detecting inhibitors of the protease activity, the helicase activity and the ATPase activity of NS3. The covalent NS4A-NS3 complexes are more soluble, stable and active than the non-covalent protease-peptide complexes previously available.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  62 CINCVCWTVYHGAGTRIIASPKGPVTQMYTNVDKDLVGWQAPQGSRSLIPCTCGSSDLYL 121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    VIRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVCTRGVAKA 181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                81
                                                                                                                                                                                                                                                                                                                                                                                                       5 GSVVIVGRINLSGD---TAYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVSTATQTFLAT 61
                                                                                                                                                                                                                                                                                                                                                                                                                           Gaps
                                                                                                       The present invention describes a covalent hepatitis C virus (HCV)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            HCV; hepatitis C virus; single chain recombinant complex; linker; NS3 protease; NS4A cofactor; NS4A-NS3 complex; serine protease; hydrophobic domain; covalent complex; detection; inhibitor.
                                                                                                                                                                                                                                                                                                                                                                         <u>ښ</u>
                                                                                                                                                                                                                                                                                                                                       87.6%: Score 904.5; DB 20; Length 216;
86.7%: Pred. No. 2.2e-87;
Live 16; Mismatches 7; Indels 3;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Yao N;
                                          New hepatitis C virus covalent complexes
                                                                          Claim 6; Page 76-77; 211pp; English
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    PC,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       AAY24945 standard; Protein; 665 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                HCV NS4A-NS3 complex SEQ ID NO:16.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             VDFIPVESLETIMRS 196
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   98US-0094331.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    98WO-US24528
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (first entry)
                                                                                                                                                                                                                                                                                                                                                         Best Local Similarity 86.7
Matches 169; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Taremi SS,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (SCHE ) SCHERING CORP.
         WPI; 1999-385385/32
                                                                                                                                                                                                                                                                                                             216 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        WO9928482-A2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   28-JUL-1998;
28-NOV-1997;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    07-SEP-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    24-NOV-1998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Malcolm BA,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      10-JUN-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Synthetic.
                                                                                                                                                                                                                                                                                                              Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     AAY24945;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               122
                                                                                                                                                                                                                                                                                                                                            Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           RESULT 14
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            AAY 24945
```

a ç 셤 ð S C

```
The present invention describes a covalent hepatitis C virus (HCV) NS4A NS3 complex comprising a central hydrophobic domain of native HCV NS4A PST in and an HCV NS3 serine protease domain. Where the hydrophobic domain of native HCV NS4A peptide is tethered by the linker to the amino terminus of the HCV NS3 protease domain. The present complex. The covalent NS4A NS3 complexes are useful for structural determination and determination of mode of binding of HCV inhibitors by NMR spectroscopy. They can also be used for detecting inhibitors of the protease activity, the helicase activity and the ATPase activity of NS3. The covalent NS4A-NS3 complexes are more soluble, stable and activity of NS3. The covalent proteases are more soluble, stable and active than the non-covalent protease-peptide complexes previously available.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 62 CINGUCWTVYHGAGTRTIASPKGPVTQMYTNVDKDLVGWQAPQGSRSLTPCTCGSSDLYL 121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       122 VTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVCTRGVAKA 181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    22 GSVVIVGRIILSGSGSITAXSQQTRGLLGCKITSLTGRDKNQVEGEVQVVSTATQSFLAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               5 GSVVIVGRINLSGD---TAYAQQTRGEEGCQETSQTGRDKNOVEGEVQIVSTATQTFLAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                HCV; hepatitis C virus; single chain recombinant complex; linker; NS3 protease; NS4a cofactor; NS4a-NS3 complex; serine protease; hydrophobic domain; covalent complex; detection; inhibitor.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  .,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              87.6%; Score 904.5; DB 20; Length 665; 86.2%; Pred. No. 1e-86; tive 16; Mismatches 8; Indels 3;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Yao N;
New hepatitis C virus covalent complexes
```

```
The present invention describes a covalent hepatitis C virus (HCV) NS4A-NS3 complex comprising a central hydrophobic domain of native HCV NS4A serine protease domain, where the hydrophobic domain of native HCV NS3 serine protease domain, where the hydrophobic domain of native HCV NS4A peptide is tethered by the linker to the amino terminus of the HCV NS3 protease domain. The present sequence represents a specifically claimed example of the above complex. The covalent NS4A NS3 complexes are useful for structural determination and determination of mode of binding of HCV inhibitors by NMR spectroscopy. They can also be used for detecting inhibitors of the protease activity, the helicase activity and the ATPase activity of NS3. The covalent NS4A NS3 complexes are more soluble, stable and active than the non-covalent protease-peptide complexes previously available.
                          New hepatitis C virus covalent complexes
                                                                                Claim 6; Page 97-99; 211pp; English.
```

665 AA; Sequence

3; Gaps Length 665; 8; Indels 87.6%; Score 904.5; DB 20; 86.2%; Pred. No. 1e-86; tive 16; Mismatches 8; : Query Match
Best Local Similarity 86.2%
Matches 169; Conservative

GSVVIVGRINLSGD---TAYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVSTATQTFLAT 61 22 GSVVIVGRIILSGSGSITAYSQQTRGLLGCIKTSLTGRDKNQVEGEVQVVSTATQSFLAT 81

62 CINGVCWTVYHGAGTRTIASPKGPVTQMYTNVDKDLVGWQAPQGSRSLTPCTCGSSDLYL 121

122 VTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVCTRGVAKA 181

Q

qq ò

ò

Ω ð

á

182 VDFIPVESLETTMRSP 197

111:1111:111111 202 VDFVPVESMETTMRSP 217

Search completed: August 30, 2003, 19:12:23 Job time : 44.6227 secs

```
GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.
```

OM protein - protein search, using sw model

August 30, 2003, 18:01:52 ; Search time 9.75674 Seconds (without alignments) 949.524 Million cell updates/sec Run on:

US-09-965-594-16 1032 1 MKKKGSVVIVGRINLSGDTA.....VAKAVDFIPVESLETIMRSP 197

Title: Perfect score: Sequence:

BLOSUM62 Gapop 10.0 , Gapext 0.5 Scoring table:

127863 seqs, 47026705 residues Searched: Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0 Maximum DB seq length: 200000000

Post-processing: Minimum Match 0% Maximum Match 100% Listing first 45 summaries

SwissProt\_41:\* Database : Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

		đ			SUMMARIES	
Result	Score	Query Match	Length	DB	ID	Description
1	884.5	85.7	3011	-	POLG_HCV1	P26664 h qenome po
7	878.5	85.1	3011	-	POLG_HCVH	enome
m	867.5	84.1	3010		POLG_HCVTW	h genome
4	857.5	ë.	3010	-	POLG_HCVJT	h genome
2	853.5	82.7	3010	٦,	POLG_HCVBK	h genome
9	853.5	ς.	3010	-	POLG_HCVJA	h genome
7	677	δ.	3033	~	POI.G_HCVJ6	h genome
60	675	5.	3033	Н	POLG_HCVJ8	h genome
6	98		321	٦	HHOA_ARATH	arabidops
10	84.5	8.2	485	Н	Y136_TREPA	
11	81		452	Н	AAMP_HUMAN	
12	78.5	7.6		Н	DEG1_ARATH	arabi
13	78			٦	PTPO_MOUSE	
14	76.5	7.4		-	CTRL_HUMAN	
15	9	7.4	323	H	VPRT_SMRVH	squir
16	76.5	7.4	333	П	MOSA_RHIME	
17	~	٠	911	Н	TB11_NEIMB	Q09056 neisseria m
18		•	503	-	PAAD_PSEAE	psend
		7.3	2663	_	CENE_HUMAN	homod
20	75	7.3	388	-	ODPT_HUMAN	P29803 homo sapien
21	75	٠	455	٦	TMS5_MOUSE	mus m
22	75	7.3	594	-	NIR_SPIOL	P05314 spinacia ol
23		٠	706	-	TRFE_HORSE	
24	74.5	7.2	764	П	ICCR_DROME	Q08180 drosophila
25	74	7.2	844	-	CN4A_RAT	
26	73.5	7.1	263	-	GRAK_MOUSE	035205 mus musculu
27	73	•		-	IBP1_HUMAN	~
28	72.5			-	MLTD_ECOLI	-
29		٠		-	THYG_RAT	a
30	72	7.0	349	-	TRPD_PSEPU	P20575 pseudomonas
31	72	•		П	GAL1_STRCO	m
32	72	7.0		-	POL_GALV	
33	72	7.0	1210	-	EGFR_MOUSE	•

P11033 mus musculu 090627 gallus gall								Q9hwc9 pseudomonas	
GRAD_MOUSE	GRAM_HUMAN	CWG2_SCHPO	PANE_RHILO	HELS_METMA	TBP1_NEIGO	MM03_RABIT	ENV_FLVGL	RPOC_PSEAE	GAL1_STRLI
7,	-			-	Н	7	,	Н	1
248	257	355	326	730	915	478	642	1399	397
6.9	6.9	0.0	6.9	6.9	6.9	8.9	8.9	8.9	8.9
71.5	71.5	3.17	77	7.1	71	70.5	70.5	70.5	70
34	36	37	9 6	40	41	42	43	44	<b>4</b>

## ALIGNMENTS

RESULT 1 10 POLGA HCV1 110 POLGA HCV1 110 POLGA HCV1 110 POLGA HCV1 111 POLGA HCV
--

```
CELLULAR AMINOPEPTIDASE.
CARSID PROTEIN ( POTENTIAL).
MATRIX PROTEIN (POTENTIAL).
MAJOR ENVELOPE PROTEIN E (POTENTIAL).
NONSTRUCTURAL PROTEIN NS1/E2 (POTENTIAL).
PROTEASE/HELICASE NS3 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4A (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4A (POTENTIAL).
                                                                                                                                                                                                                                                                                                    RNA-DIRECTED RNA POLYMERASE (POTENTIAL) POTENTIAL.
                                                                                                                                                                                           Polyprotein; Glycoprotein; Transferase; RNA-directed RNA polymerase;
Core protein; Coat protein; Envelope protein; Helicase; ATP-binding;
Transmembrane; Nonstructural protein; Hydrolase; Serine protease;
                                                                                                                                                                                                                                                                                                                   SIMILARITY).
SIMILARITY).
SIMILARITY).
                                                                                                                                                                                                                        REMOVED FROM CAPSID PROTEIN C BY THE
                                                                                                                                                                                                                                                                                                                                                                      (POTENTIAL).
(POTENTIAL).
(POTENTIAL).
(POTENTIAL).
(POTENTIAL).
(POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                    (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                            (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                           (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                   (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                          (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                            (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Length 3011;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               MW; 65F8C9447FCE5AF9 CRC64;
                                                                                                                                                                                                                                                                                                                   CHARGE RELAY SYSTEM (BY
CHARGE RELAY SYSTEM (BY
CHARGE RELAY SYSTEM (BY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Score 884.5; DB 1;
Pred. No. 4.3e-76;
                                                                                                                                                                                                                                                                                                                                                                                                                                   (GLCNAC.
(GLCNAC.
(GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                              (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                   (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                          (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                       (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                 (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                              (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                         (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                          (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          GLCNAC
                                                                                                                                                                                                                                                                                                                                         (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                N-LINKED
                                                                                                                                                                                                                                                                                                                                                          N-LINKED
                                                                                                                                                                                                                                                                                                                                                  DECH BOX
           85.7%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               327197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  645
2041
2077
2240
2364
2789
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  645
2041
2077
2240
2364
2789
3011
                                                                                                                                                                                                                  3D-structure.
INIT_MET
                                                                                                                                                                                                                                                                                                                                                                                                            CARBOHYD
CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                   ACT_SITE
ACT_SITE
ACT_SITE
NP_BIND
SITE
                                                                                                                                                                                                                                                                                                                                                                                       CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Query Match
Best Local S
                                                                                                                                                                                                                                                                                                     CHAIN
TRANSMEM
                                                                                                                                                                                                                                                                                                                                                                                                      SARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                     CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                             CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                   CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                          CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SEQUENCE
                                                                                                                                                                                                                                                                                                                                                          CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                 CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                        CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                               CARBOHYD
                                                                                                                                                                                                                                               CHAIN
                                                                                                                                                                                                                                                                      CHAIN
CHAIN
CHAIN
                                                                                                                                                                                                                                                              CHAIN
                                                                                                                                                                                                                                                                                              CHAIN
```

```
1125 GGSSDLYLVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAV 1184
                                                                                                                                   54 ATQTFLATCINGVCWTVYHGAGTRTIASPKGPVTQMYTNVDKDLVGWQAPQGSRSLTPCT 113
                                                                                                                                                                                                                                114 CGSSDLYLVTRHADVIPVRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAV 173
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             01-40G-1992 (Rel. 23, Created)
01-40G-1992 (Rel. 23, Last sequence update)
01-52EP-2003 (Rel. 42, Last annotation update)
01-60C-1992 (Rel. 42, Last annotation update)
03-60C-1992 (Rel. 42, Last annotation update)
04-60C-1992 (GP63) (GP32) (GP32) (GP63) (GP68) (GP06) (NS1): Protein P7; Nonstructural protein NS2 (P21)
05-73 (GP68) (GP70) (NS1): Protease/helicase NS3 (P70) (Hepacivitin)
05-73 (P77); Nonstructural protein NS4 (P4); Nonstructural protein
05-74 (P77); Nonstructural protein NS5 (P66); Nonstructural protein
05-75 (P77); Nonstructural Protein NS5 (P66) (P77) (P
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS) OF 1192-1657.

WEDLINE-9815421; PubMed-9493270;

Kim J.L., Morgenstern K.A., Griffith J.P., Dwyer M.D., Thomson J.A.,

MUTCKO M.A., Lin C., Caron P.R.;

WHEPPLINE WITH STAND TO C., Caron P.R.;

WITHOUS THE C. CARON P.R.;

SHOOMULE OF THE CIPSTAL STRUCTURE PROVIDES INSIGHTS Into the mode of unwinding.";

STUTUCTURE 6 899-100(1998).

STUTUCTURE SEPPONSIBLE FOR THE CLEAVAGE OF NS2-NS3.

--- FUNCTION: PROTEASE NS3 IS RESPONSIBLE FOR THE CLEAVAGE OF
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                MEDLINE-97331322; PubMed-9187654;
Yao N., Hesson T., Cable M., Hong Z., Kwong A.D., Le H.V., Weber P.C.;
"Structure of the hepatitis C virus RNA helicase domain.";
Nat. Struct. Biol. 4:465-467(1997).
ACTIVATION OF NGS.

-!- FUNCTION: NS54 SEEMS TO HAVE A TRANSCRIPTIONAL ACTIVATORY ROLE.
-!- FUNCTION: NS58 IS A RNA-DEPENDENT RNA POLYMERASE THAT PLAYS AN ESSENTIAL ROLE IN THE VIRUS REPLICATION.
-!- CATALYTIC ACTIVITY: Hydrolysis of four peptide bonds in the viral precursor polyprotein, commonly with Asp or Glu in the P6 position, Cys or Thr in P1 and Ser or Ala in P1'.
-!- CATALYTIC ACTIVITY: N nucleoside triphosphate - N diphosphate +
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SUBGNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS: E1 AND E2. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MRNA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            NS3-NS4A, NS4A-NS4B, NS4B-NS5A AND NS5A-NS5B.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    U
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             "Genomic structure of the human prototype strain H of hepatitis virus: comparison with American and Japanese isolates."; Proc. Natl. Acad. Sci. U.S.A. 88:10292-10296(1991).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                MEDLINE-92052256; PubMed-1658800;
Inchauspe G., Zebedee S., Lee D.H.H., Sugitani M., Nasoff M.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    X-RAY CRYSTALLOGRAPHY (2.1 ANGSTROMS) OF 1207-1657.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                PRT; 3011 AA
                                                                                                                                                                                                                                                                                                                                                                                               174 CTRGVAKAVDFIPVESLETTMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SEQUENCE FROM N.A. MEDLINE-92052256;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Hepacivirus.
NCBI_TaxID=11108;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Prince A.M.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             POLG_HCVH
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        RESULT 2
POLG_HCVH
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               q
                                                                                                                                                              d
                                                                                                                                                                                                                                   õ
                                                                                                                                                                                                                                                                                       g
```

9; Mismatches

Conservative

Similarity

Local Sim

Matches

9: Gaps

Indels

 $\sim$ 

```
POTENTIAL)
(POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Length 3011;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     327142 MW; 772CBB29CCD94753 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Score 878.5; DB 1;
Pred. No. 1.6e-75;
                            (GLCNAC.
(GLCNAC.
(GLCNAC.
                                                                                                               (GLCNAC.
(GLCNAC.
(GLCNAC.
                          N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            85.1%;
83,3%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ¥,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1584
1598
1606
1614
1622
1627
1635
1640
3011
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1480
1481
1489
1497
1507
                            STRAND
STRAND
STRAND
HELIX
SEQUENCE
          CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    STRAND
HELIX
STRAND
HELIX
TURN
STRAND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    STRAND
STRAND
STRAND
STRAND
STRAND
STRAND
STRAND
STRAND
STRAND
HELIX
HEL
        This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation-the Buropean Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SMART; SM00487; DEXDC; 1.
Polyprotein; Glycoprotein; Transferase; RNA-directed RNA polymerase; Core protein; Coat protein; Envelope protein; Helicase; ATP-binding; Transmembrane; Nonstructural protein; Hydrolase; Serine protease;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            SIMILARITY).
SIMILARITY).
SIMILARITY).
PIM: THE STRUCTURAL PROTEINS C, EL AND EZ ARE PRODUCED BY PROTEOLYTIC PROCESSING BY THE HOST SIGNAL PEPTIDASES. SIMILARITY: THE NS2 PROTEASE BELONGS TO PEPTIDASE FAMILY U39. SIMILARITY: THE NS3 PROTEASE BELONGS TO PEPTIDASE FAMILY S29.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                N-LINKED (GLCNAC. . .) (POTENTIAL).
N-LINKED (GLCNAC. . .) (POTENTIAL).
N-LINKED (GLCNAC. . .) (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      REMOVED FROM CAPSID PROTEIN C
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ENVELOPE GLYCOPROTEIN E1.
ENVELOPE GLYCOPROTEIN E2.
PROTEIN P7.
PROTEASE/HELICASE NS3.
NONSTRUCTURAL PROTEIN NS2.
NONSTRUCTURAL PROTEIN NS4A.
NONSTRUCTURAL PROTEIN NS4B.
NONSTRUCTURAL PROTEIN NS5B.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              (BY
(BY
(BY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CELLULAR AMINOPEPTIDASE. CAPSID PROTEIN C.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CHARGE RELAY SYSTEM
CHARGE RELAY SYSTEM
CHARGE RELAY SYSTEM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ATP (POTENTIAL).
DECH BOX.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            InterPro; IPRO01809; NCVNSSA; InterPro; IPR001808; NCVNSSA; InterPro; IPR001808; NCVNSSA; InterPro; IPR001809; NALPOLLSSA; InterPro; IPR001809; NALPOLLSSA; InterPro; IPR007099; NALPOLLSSA; InterPro; IPR007099; NALPOLLSSA; InterPro; IPR007099; NALPOLLSSA; NCVCORE; InterPro; IPR001809; NCVCORE; InterPro; IPR001809; NCVCORE; Interpro; IPR001809; NCVNSS; Interpro; IPR001809; NCVNSS; Interpro; IPR001809; NCVNSS; Interpro; IPR001809; NCVNSSA; Interpro; IPR00
                                                                                                                                                                                                                                                                                                                                                                                                                                         InterPro; IPR001410; DEAD.
InterPro; IPR002522; HCV_capsid.
InterPro; IPR002521; HCV_core.
InterPro; IPR002531; HCV_RO;
InterPro; IPR002531; HCV_NS2.
InterPro; IPR004109; HCV_NS2.
InterPro; IPR004109; HCV_NS3.
InterPro; IPR000145; HCV_NS3.
InterPro; IPR001490; HCV_NS4.
InterPro; IPR001490; HCV_NS4.
                                                                                                                                                                                                                                                                            EMBL; M67463; AAA45534.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1026
1657
1711
1972
2420
3011
369
1107
11165
11165
1237
1319
1319
1319
209
                                                                                                                                                                                                                                                                                                                 PDB: 1HEI; 25-NOV-98.
PDB: 1A1V: 16-FEB-99.
PDB: 1A1R: 17-JUN-98.
MEROPS: $29.001; -.
                                                                                                                                                                                                                                                                                                    A36814; GNWVCH.
                                                                                                                                                                                                                                                                                                                                                                                                                     TRANSFAC; TO4155;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               3D-structure.
INIT_MET
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CARBOHYD
CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ProDom; E
SMART; SN
```

```
CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Ouery
   This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bloinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/
                                                                                                       114 CGSSDLYLVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAV 173
                                                                                                                                                                                                                   01-APR-1993 (Rel. 25, Created)
01-APR-1993 (Rel. 25, Last sequence update)
115-SEP-2003 (Rel. 42, Last aenochation update)
Genome polyprotein (Contains Capsid protein C (Core protein) (P22);
Envelope glycoprotein El (GP32) (GP35); Envelope glycoprotein E2
(GP68) (GP70) (N31); Protein P7; NonStructural protein NS2 (P21)
(EC 3.4.22.-1; Protease/helicase NS3 (P70) (Hepacivirin)
(EC 3.4.21.98); NonStructural protein NS54 (P65); NonStructural protein NS58 (P65) (P77); NonStructural protein NS58 (P65) (P70) (RNA-directed RNA POlymerase) (EC 2.7.7.48)].
Hepatitis C virus (Isolate Taiwan) (HCV).
Viruses; SERNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                             3 KKGSVVIVGRIN-----LSGDTAYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVST
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        (RNA)(N).
SUBONIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A
  Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS: PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MRNA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY S29.
15; Indels
                                                                                                                                                                                                   PRT; 3010 AA
 10; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  or send an email to license@isb-sib.ch).
                                                                                                                                             174 CTRGVAKAVDFIPVESLETTMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     EMBL; M84754; -; NOT_ANNOTATED_CDS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              InterPro; IPR001410; DEAD
 Matches 170; Conservative
                                                                                                                                                                                                    STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       PDB; 1N64; 25-FEB-03.
PDB; 1NS3; 08-APR-98
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              PIR; A40244; GNWVTW
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           MEROPS; S29.001; ..
                                                                                                                                                                                                                                                                                                                                              NCBI_TaxID=31645;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     U39.001
                                                                                                                                                                                                                                                                                                                                     Hepacivirus
                                                                                                                                                                                                    POLG_HCVTW
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     MEROPS;
                                                                                                                                                                                                    a
                                                          å
                                                                         g
                                                                                             ò
                                                                                                               셤
                                                                                                                                 ò
                                                                                                                                                  d
```

```
CELLULAR AMINOPEPTIDASE.

CELLULAR AMINOPEPTIDASE.

MATRIX PROTEIN (POTEBTIAL).

MAJOR ENVELOPE PROTEIN E (POTENTIAL).

MAJOR ENVELOPE PROTEIN BOUSTROCTURAL PROTEIN NSI, FE OPTEMITAL).

NONSTRUCTURAL PROTEIN NS2 (POTEMITAL).

PROTEASE/HELICASE NS3 (POTEMITAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      NONSTRUCTURAL PROTEIN NS4A (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4B (POTENTIAL).
RNA-DIRECTED RNA POLYMERASE (POTENTIAL)
POTENTIAL.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Core protein; Coat protein; Envelope protein; Helicase; ATP-binding;
Transmembrane; Nonstructural protein; Hydrolase; Serine protease;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Polyprotein; Glycoprotein; Transferase; RNA-directed RNA polymerase
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               REMOVED FROM CAPSID PROTEIN C BY THE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SIMILARITY).
SIMILARITY).
SIMILARITY).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Ġ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          MW; AAD267D55CDFE215 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CHARGE RELAY SYSTEM (BY
CHARGE RELAY SYSTEM (BY
CHARGE RELAY SYSTEM (BY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    GLCNAC. . .)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (GLCNAC. . . . (GLCNAC. . . .
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         GLCNAC. . .
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                DB 1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    N-LINKED (GLCNAC,
N-LINKED (GL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Score 867.5; DB 1
Pred. No. 1.8e-74;
8; Mismatches 14
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ATP (POTENTIAL).
DECH BOX.
Pro; IPRO000/40,

Pro; IPRO01469; HCV.NS4D.

PPro; IPRO01669; HCV.NS5A.

EPPro; IPRO010669; HCV.NS5A.

EPPro; IPRO01065; HCV.AGRP.

EPPro; IPRO07094; RNA_pol_DS_PS.

EPPro; IPRO07094; RNA_pol_PS_PS.

EMPRO143; HCV_capsid. 1.

Emp. PFO1143; HCV_capsid. 1.

Fam; PFO1143; HCV_capsid. 1.

Fam; PFO11539; HCV_cnv. 1.

Fam; PFO11539; HCV_NS1; 1.

Pfam; PFO11509; HCV_NS3; 1.

Pfam; PFO11509; HCV_NS4; 1.

Pfam; PFO11006; HCV_NS4s; 1.

Pfam; PFO11001; HCV_NS4s; 
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  18:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                84.1%;
79.9%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              327047
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 115
191
383
729
1006
1615
1862
2013
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 3D-structure.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Best Local Sim
Matches 163;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CARBOHYD
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   TRANSMEM
ACT_SITE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ACT_SITE
NP_BIND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       PARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CARBOHYD
```

ü

S

```
Best Local Similarity 78.4 Matches 160; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               3 KKGSVVIVGRIN---
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Similarity
                   InterPro;
InterPro;
InterPro;
InterPro;
InterPro;
InterPro;
   InterPro;
                                                                              (nterPro;
                                                                                       InterPro;
                                                                                                  InterPro;
                                                                                                           interPro;
                                                                                                                                                                                                                                                                                                                                                                                                     ACT_SITE
ACT_SITE
ACT_SITE
NP_BIND
                                                                                                                                                                                                                                                                                                                                                                                  CHAIN
IRANSMEM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Query Match
                                                                                                                                                                                                                           ProDom;
SMART; S
                                                                                                                                                                                                                                                                                                                                                                                                                                                     CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                               CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CARBOHYD
                                                                                                                                                                                                                                                                                    INIT_MET
                                                                                                                                                          Pfam;
                                                                                                                                                                                                                                                                                                        CHAIN
  ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       This SWISS-PROT entry is copyright. It is produced through a collaboration
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce.
                                                                                      1005 RRGREILLGPADSLEGRGWRLLAPITAYAQQTRGLFGCIITSLTGRDKNQVEGEVQVVST 1064
                                                  54 ATQTFLATCINGVCWIVYHGAGTRTIASPKGPVTQMYTNVDKDLVGWQAPQGSRSLTPCI 113
                                                                             114 CGSSDLYLVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAV 173
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 the Swiss Institute of Bioinformatics and the EMBL outstation
3 KKGSVVIVGRIN------LSGDTAYAQQTRGEEGCQETSQTGRDKNOVEGEVQIVST 53
                                                                                                                                                                                                      01-APR-1993 (Rel. 25, Created)
01-APR-1993 (Rel. 25, Last sequence update)
15-SEP-2003 (Rel. 42, Last amonotation update)
Genome polyprotein (Contains Capsid protein C (Core protein) (P22);
Envelope glycoprotein El (GP32) (GP35); Envelope glycoprotein E2
(GP68) (GP70) (NS1); Protein P7; Nonstructural protein NS2 (P21)
(EC 3.4.21.98); Nonstructural protein NS48 (P4); Nonstructural protein
NS48 (P27); Nonstructural protein NS5A (P4); Nonstructural protein
NS5B (P60) (RNA-directed RNA polymerase) (EC 2.7.7.48)].
Hepatitis C virus (isolate HC-JT) (HCV).
                                                                                                                                                                                                                                                                                                                                                                      SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A LIPOPRACTEIN ENVELOPE. CONSISTS OF TWO PROTEINS: PROFEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MRNA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY S29.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                or send an email to license@isb-sib.ch)
                                                                                                                             174 CTRGVAKAVDFIPVESLETTMRSP 197
                                                                                                                                                                                                                                                                                                                                                     SEQUENCE FROM N.A.
MEDLINE-92295714; PubMed-1318627;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    EMBL; D11168; BAA01943.1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             InterPro; IPR001410; DEAD.
                                                                                                                                                                                      STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       PDB; 1A1Q; 25-WAR-98.
PDB; 1JXP; 14-JAN-98.
WEROPS; S29.001; -.
WEROPS; U39.001; -.
                                                                                                                                                                                                                                                                                                                                   MCBI_TaxID=31642;
                                                                                                                                                                                    POLG_HCVJT
Q00269;
                                                                                                                                                                                       g
                                       ò
                                                        g
                                                                            ò
                                                                                              g
                                                                                                                   οy
                                                                                                                                   셤
```

```
CARSID PROTEIN C (POTENTIAL).
MATRIX PROTEIN (POTENTIAL).
MAJOR ENVELOPE PROTEIN E (POTENTIAL).
NONSTRUCTURAL PROTEIN NSJ (POTENTIAL).
PROTEASE/HELICASE NS3 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS2 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4A (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4B (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Polyprotein; Glycoprotein; Transferase; RNA-directed RNA polymerase; Core protein; Coat protein; Envelope protein; Helicase; ATP-binding; Transmembrane; Nonstructural protein; Hydrolase; Serine protease;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              RNA-DIRECTED RNA POLYMERASE (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   REMOVED FROM CAPSID PROTEIN C BY THE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SIMILARITY).
SIMILARITY).
SIMILARITY).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       (POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      POTENT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             94A1C77435D642BB CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CHARGE RELAY SYSTEM (BY
CHARGE RELAY SYSTEM (BY
CHARGE RELAY SYSTEM (BY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CELLULAR AMINOPEPTIDASE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ATP (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   N-LINKED (
N-LINKED (
N-LINKED (
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      N-LINKED
N-L
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     DECH BOX
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       N-LINKED
                                                                                                                                                                                                                                                                                   IPR002868; HCV_NS5a.
IPR002166; HCV_RdRP.
IPR007095; RNA_pol_DS_PS.
IPR007094; RNA_pol_PSvir.
                                 HCV_core.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                PF02907; HCV_NS3; 1.
PF01006; HCV_NS4s; 1.
PF01506; HCV_NS5s; 1.
PF0257; helicase_C; 1.
PF00298; Viral_RGRP; 1.
                                                                                                                                                                                                                                             HCV_NS4b
                                                                                                                                                                                                                                                                                                                                                                                                                           Pfam; PF01543; HCV_capsid; 1.
Pfam; PF01542; HCV_core; 1.
Pfam; PF01539; HCV_env; 1.
Pfam; PF01560; HCV_NS1; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             326573
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  HCV_NS2;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             DEXDC:
IPR002522;
                                                                                                                                                                             IPR004109;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ¥.
                                                                                                     PR002531;
                                                                                                                                                                                                                                                 IPR001490;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     3010
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            PF01560; PF01538;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SM00487;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Pfam; PF02907;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Pfam; PF01006;
Pfam; PF01001;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Pfam; PF01506;
Pfam; PF00271;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Pfam; PF00998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 3D-structure.
```

Ä 53

9; Gaps

15; Indels

20; Mismatches Score 857.5;

DB 1; Length 3010;

1.6e-73

Pred.

83.1%; 78.4%;

---LSGDTAYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVST

```
1065 ATQSFLATCVNGVCWTVFHGAGSKTLAGPKGPITOMYTNVDQDLVGWHAPPGARSLTPCT 1124
                                                                                                                                                                                                     RRGREILLGPADSIEGQGWRLLAPITAYAQQTRGLLGCIVTSLTGRDKNQVEGEVQVVST 1064
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   X. ARY CRYSTALLOGRAPHY (2.2 ANGSTROMS) OF 1027-1210 AND 1678-1691.
X. MEDLINE-88227846; PubMed-9568891;
YAN Y., Li Y., Munshi S., Sardana W.,
Yan Y., Li Y., Munshi S., Sardana W.,
Steinkuehler C., Tomei L., de Francesco R., Kuo L.C., Chen Z.;
Steinkuehler C., Tomei L., de Francesco R., Kuo L.C., Chen Z.;
T. Complex of NS3 protease and NS4A peptide of BK strain hepatitis C.
T. VILUS: a 2.2-A resolution structure in a hexagonal crystal form.";
Protein Sci. 7:837-847 (1998)
C. I. FUNCTION: THE SMALL PROTEINS NS2A, NS2B, NS4A AND NS4B ARE
HYDROPHOBIC, SUGGESTING A POSSIBLE MEMBRANE-REIATED FUNCTION.
NS3 AND NS5 NAY PLAY A ROLE IN THE VIRAL RNA REPLICATION.
NS3 AND NS5 NAY PLAY A ROLE IN THE VIRAL RNA REPLICATION.
C. I- CATALTIC ACTIVITY: Hydrolysis of four peptide bonds in the P6
prostition, Cys or Thr in Pl and Ser or Ala in Pl'.
C. CATALTIC ACTIVITY: N nucleoside triphosphate - N diphosphate +
                                                                          54 ATQTFLATCINGVCWTVYHGAGTRTIASPKGPVTQWYTNVDKDLVGWQAPQGSRSLTPCT 113
                                                                                                                                                                          CGSSDLYLVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAV 173
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               01-AUG-1992 (Rel. 23, Created)
01-AUG-1992 (Rel. 23, Last sequence update)
115-SEP-2003 (Rel. 24, Last annocation update)
16-SEP-2003 (Rel. 42, Last annocation update)
Genome polyprotein (Contains: Capsid protein C (Core protein) (P22);
Envelope glycoprotein El (GP32) (GP35); Envelope glycoprotein E2
(GP68) (GP70) (NS1); Protein P7: Nonstructural protein NS2 (P20)
(EC 3.4.22.-); Protease/helicase NS3 (P70) (Hepacivirin)
(EC 3.4.21.98); Nonstructural protein NS4A (P4); Nonstructural protein NS4B (P27); Nonstructural protein NS5A (P56); Nonstructural protein NS4B (P66) (P70) (RNA-directed RNA polymerase) (EC 2.7.7.48)].
Hepatitis C virus (isolate BNK) (RVA)
Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Takamizawa A., Mori C., Fuke I., Manabe S., Murakami S., Fujita J., Onishi E., Andoh T., Yoshida I., Okayama H.; Structure and organization of the hepatitis C virus genome isolated
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 X-RAY CRYSTALLOGRAPHY (2.4 ANGSTROMS) OF 1027-1215.

MEDLINE-97015088; PubMed-8861916;
LOVE R.A., Parge H.E., Mickersham J.A., Hostomsky Z., Habuka N.,
MOGNAW E.W., Adachi T., Hostomska Z.;
MThe crystal structure of hepatitis C virus NS3 proteinase reveals
trypsin-like fold and a structural zinc binding site.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   MEDLINE-96235224; PubMed-8647104;
Borowski P., Heiland M., Oehlmann K., Becker B., Korneteky L.;
"Non-structural protein 3 of hepatitis C virus inhibits
phosphorylation mediated by cAMP-dependent protein kinase.";
Eur. J. Biochem. 237:611-618(1996).
                                                                                                                                                                                                                                                                                                                                                                                                                                                       PRT; 3010 AA
                                                                                                                                                                                                                                                                                                          1185 CTRGVAKAVDFIPVESMETTMRSP 1208
                                                                                                                                                                                                                                                                               174 CTRGVAKAVDFIPVESLETTMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SEQUENCE FROM N.A. MEDLINE-91140698; PubMed-1847440;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Virol. 65:1105-1113(1991).
                                                                                                                                                                                                                                                                                                                                                                                                                                                          STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SEQUENCE OF 1487-1500
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    from human carriers.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Hepacivirus,
NCBI_TaxID=11105;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           { RNA } (N).
                                                                                                                                                                                                                                                                                                                                                                                                                                                       POLG_HCVBK
P26663;
                                                                                                                                                                               114
                                                                                                                                                                                                                                                                                                                                                                                                          RESULT 5
POLG_HCVBK
                                                                                                                                                                                                                                                                                                                                                                                                                                                            ALD PARTER TO THE STATE OF THE 
                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                            ç
                                                                             à
                                                                                                                                                                               ò
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its most by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@lsb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CELLULAR ANINOPEPTIDASE.
CAPSID PROTEIN (POTENTIAL).
MAJOR ENVELOR (POTENTIAL).
MAJOR ENVELOR PROTEIN E (POTENTIAL).
NONSTRUCTURAL PROTEIN NSI/F2 (POTENTIAL).
PROTEINEAE/HELICASE NSI (POTENTIAL).
PROTEINEAE/HELICASE NSI (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4A (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4A (POTENTIAL).
RNA-DIRECTED RNA POLYMERASE (POTENTIAL).
SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Polyprotein; Glycoprotein; Transferase; RNA-directed RNA polymerase;
Core protein; Coat protein; Envelope protein; Helicase; ATP-binding;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Core protein; Coat protein; Envelope protein; Holicase; ATP-binding;
Transmembrane; Nonstructural protein; Hydrolase; Serine protease;
                     LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS: PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MRNA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              REMOVED FROM CAPSID PROTEIN C BY
                                                                                                SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY S29.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CHARGE RELAY SYSTEM.
CHARGE RELAY SYSTEM.
CHARGE RELAY SYSTEM.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ATP (POTENTIAL)
DECH BOX.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 POTENTIAL.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           InterPro; 1979.01.; interPro; 1PR001252; HCV_capsid.
InterPro; IPR001252; HCV_capsid.
InterPro; IPR001251; HCV_care.
InterPro; IPR001251; HCV_env.
InterPro; IPR001251; HCV_NS1.
InterPro; IPR001251; HCV_NS2.
InterPro; IPR00140; HCV_NS4.
InterPro; IPR00140; HCV_NS4.
InterPro; IPR00140; HCV_NS4.
InterPro; IPR00140; HCV_NS4.
InterPro; IPR00126; HCV_NS4.
InterPro; IPR00126; HCV_NS4.
InterPro; IPR00106; HCV_NS4.
InterPro; IPR00106; HCV_RGRP.
InterPro; IPR001094; RNA_POL_PSY1r.
Pfam; PF01540; HCV_core; I.
Pfam; PF01540; HCV_core; I.
Pfam; PF01550; HCV_NS1; I.
Pfam; PF01006; HCV_NS1; I.
Pfam; PF01001; HCV_NS1; I.
Pfam; PF01001; HCV_NS3; I.
Pfam; PF01001; HCV_NS3; I.
Pfam; PF01001; HCV_NS4; I.
Pfam; PF01001; HCV_NS4; I.
Pfam; PF01001; HCV_NS4; I.
Pfam; PF01001; HCV_NS3; I.
Pfam; PF01001; HCV_NS3; I.
Pfam; PF01001; HCV_NS3; I.
Pfam; PF01001; HCV_NS3; I.
                                                                                                                                                                                                                                                                                                                                                EMBL; M58335; AAA72945.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 115
191
191
729
1006
11615
2013
3010
369
11083
11165
11165
11165
11165
11165
11165
11165
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   PDB: 1CSJ: 08-NOV-99.
PDB: 1GX5: 09-APR-02.
PDB: 1GX6: 10-APR-02.
PDB: 1QUV: 26-JUN-00.
PDB: 80HM; 20-APR-99.
                                                                                                                                                                                                                                                                                                                                                                           PIR; A38465; GNWVTC.
PDB: 1A10; 25-MAR-98
PDB: 1JXP; 14 JAN-98.
PDB: 1NS3; 08-APR-98
PDB; 1C2P; 15-NOV-00.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         116
192
384
730
1007
1616
1863
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1083
1107
1165
1230
1316
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   MEROPS; S29.001
MEROPS; U39.001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         3D-structure.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ACT_SITE
ACT_SITE
ACT_SITE
ACT_SITE
NP_BIND
SITE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 TRANSMEM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CHAIN
CHAIN
CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CHAIN
```

```
1065 ATOSFLATCVNGVCWTVYHGAGSKTLAAPKGPITOMYTNVDQDLVGWPKPPGARSLTPCT 1124
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    125 CGSSDLYLVTRHADVIPVRRRGDSRGSLLSPRPVSYLKGSSGGPLLCPFGHAVGIFRAAV 1184
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              113
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            114 CGSSDLYLYTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              3 KKGSVVIVGRIN-----LSGDTAYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVST
                                                                  (POTENTIAL)
(POTENTIAL)
(POTENTIAL)
(POTENTIAL)
(POTENTIAL)
(POTENTIAL)
(POTENTIAL)
(POTENTIAL)
                                                                                                                                                             (POTENTIAL).
(POTENTIAL).
(POTENTIAL).
                                                                                                                                                                                               (POTENTIAL).
(POTENTIAL).
(POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Length 3010;
                                                                                                                                                            Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 82.7%; Score 853.5; DB 1;
77.9%; Pred. No. 3.9e-73;
tive 21; Mismatches 15;
                                                                    (GLCNAC.
                                  (GLCNAC.
(GLCNAC.
(GLCNAC.
                                                                                                                   (GLCNAC.
                                                                                             (GLCNAC.
                                                                                                       (GLCNAC.
                                                                                                                                         (GLCNAC.
            (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Local Similarity
tes 159; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1185
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               54
                                                                                                                                                                                                  CARBOHYD
CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           174
                                   CARBOHYD
CARBOHYD
                                                                                             CARBOHYD
CARBOHYD
                                                                                                                              CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Query Match
                                                            CARBOHYD
                                                                      CARBOHYD
                                                                                                                    CARBOHYD
                                                                                                                                                                  CARBOHYD
                                                                                                                                                                             CARBOHYD
                                                                                                                                                                                        CARBOHYD
                                                                                                                                                                                                                                                                                                                                        STRAND
STRAND
STRAND
TURN
STRAND
STRAND
STRAND
STRAND
                                                                                                                                                                                                                                                                                                                                                                                                                                                          STRAND
STRAND
HELIX
TURN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      TURN
STRAND
TURN
STRAND
                                                                                                                                                                                                                                                         STRAND
STRAND
STRAND
TURN
                                                                                                                                                                                                                                                                                                       STRAND
HELIX
TURN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 STRAND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               STRAND
                                                                                                                                                                                                                                     STRAND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Best Loca
Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          HELIX
STITE 
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            οŽ
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its way non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                              01-A0G-1992 (Rel. 23, Created)
01-A0G-1992 (Rel. 23, Last sequence update)
01-A0G-1992 (Rel. 41, Last annotation update)
28-FEB-2003 (Rel. 41, Last annotation update)
Genome polyprotein [Contains: Capsid protein C (Core protein) (P22);
Envelope glycoprotein E1 (GP32) (GP35); Envelope glycoprotein E2 (GP68) (GP70) (NS1); Protein P7; Nonstructural protein NS2 (P21) (RC 3 4.22.-); Proteasses/helicase NS3 (P70) (Hepaclytin)
(RC 3 4.21.98); Nonstructural protein NS5A (P56); Nonstructural protein NS5B (P66) (P70) (RNA-directed RNA polymerase) (EC 2.7.7.48)].
NS5B (P66) (P70) (RNA-directed RNA polymerase) (EC 2.7.7.48)].
VIRUSES; SSRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     MEDLINE-91192160; PubMed-1849488;
Rato N., Hijiatta M., Nakagawa M., Ootsuyama Y., Muraiso K.,
Rato N., Hijiatta M., Nakagawa M., Ootsuyama Y., Muraiso K.,
Namotohon S., Shimotohon G. the Japanese hepatitis C viral genome.";
"Molecular structure of the Japanese hepatitis C viral genome.";
"Molecular structure of the Japanese hepatitis C viral genome.";
"FEBS Lett. 280:325-328(1991).
"I- FUNCTION: THE SMALL PROTEINS NS2A, NS2B, NS4A AND NS4B ARE
HYDROPHOBIC, SUGGESTING A POSSIBLE MEMBRANE-RELATED FUNCTION.
NS3 AND NS5 MAY PLAY A ROLE IN THE VIRAL RNA REPLICATION.
-!- CATALYTIC ACTIVITY: Hydrolysis of four peptide bonds in the viral position, Cys or Thr in Pl and Ser or Ala in Pl'.
-!- CATALYTIC ACTIVITY: N nucleoside triphosphate - N diphosphate +
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY LIPOPROTEIN ENVELOPE CONSISTS OF TWO PROTEINS: PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MRNA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  MEDINE-91088550; PubMed-2175903; Kato N., Hijikata M., Ootsuyama Y., Nakagawa M., Ohkoshi S., Sugimura T., Shimotohno K.; "Molecular cloning of the human hepatitis C virus genome from Japanese patients with non-A, non-B hepatitis."; Proc. Natl. Acad. Sci. U.S.A. 87:9524-9528(1990).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY S29.
3010 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   HCV_core.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             HCV_NS3.
HCV_NS4a.
HCV_NS4b.
HCV_NS5a.
HCV_RGRP.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          HCV_env.
HCV_NS1.
HCV_NS2.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                EMBL; D90208; BAA14233.1; -. PIR; A39253; GNWVCJ.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               InterPro; IPR001410;
InterPro; IPR002522;
InterPro; IPR002521;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             InterPro; IPR002519;
InterPro; IPR002531;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 IPR004109;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             IPR002518;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       IPR002166:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             IPR001490;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   PIR; A39253; GNWVCJ
HSSP; P26663; lJXP.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               MEROPS; S29.001; -
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         NCBI_TaxID-11116;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (RNA) (N).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       InterPro;
InterPro;
       POLG_HCVJA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        InterPro;
                               P26662,
```

∞

```
ï
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1005 RRGKEILLGPADSFGEQGWRLLAPITAYSQQTRGLLGCIITSLTGRDKNQVDGEVQVLST 1064
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              54 ATQTFLATCINGVCWTVYHGAGTRTIASPKGPVTQMYTNVDKDLVGWQAPQGSRSLTPCT 113
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    3 KKGSVVIVGRINLSGD-----TAYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVST 53
                                                                                                                                                                                                                                                   CAPSID PROTEIN C (POTENTIAL).
MAIRIX PROTEIN (POTENTIAL).
MAJOR ENVELOPE PROTEIN E (POTENTIAL).
NONSTRUCTURAL PROTEIN NS1 (POTENTIAL).
NON-STRUCTURAL PROTEIN NS2 (POTENTIAL).
PROTEASE/HELICASE NS3 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS44 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS48 (POTENTIAL).
RNA-DIRECTED RNA POLYMERASE (POTENTIAL).
                                                                                                                                                                 Gaps
                                                                                                                                                                                                                                                                                                                                                                      CHARGE RELAY SYSTEM (BY SIMILARITY).
CHARGE RELAY SYSTEM (BY SIMILARITY).
CHARGE RELAY SYSTEM (BY SIMILARITY).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                               (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              6
                                                                                                                                                                                                                                                                                                                                                                                                                                  (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     82.7%; Score 853.5; DB 1; Length 3010; 77.0%; Pred. No. 3.9e-73;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CRC64:
                                                                                                                                                                                                                                                                                                                                                                                                                               N-LINKED GECNAC...)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              AA993794F46DB185
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            15;
                                                                                                                                                                                                                                                                                                                                                                                                           ATP (POTENTIAL).
DECH BOX.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            23; Mismatches
                                                                                                                                                                                                                                                                                                                                                            POTENTIAL.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    N-LINKED
InterPro; IPR001650; Helicase_C.
InterPro; IPR007095; RNA_pol_DS_PS.
InterPro; IPR007094; RNA_pol_DS_PS.
Pfam; PF01543; HCV_capsid; l.
Pfam; PF01542; HCV_core; l.
Pfam; PF01542; HCV_core; l.
Pfam; PF01559; HCV_core; l.
Pfam; PF01560; HCV_NS3; l.
Pfam; PF01560; HCV_NS3; l.
Pfam; PF01006; HCV_NS3; l.
Pfam; PF01006; HCV_NS4a; l.
Pfam; PF01506; HCV_NS4b; l.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             174 CTRGVAKAVDFIPVESLETTMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   645 645 N-LIN
2041 2041 N-LIN
2077 2077 N-LIN
2740 2240 N-LIN
2788 2788 N-LIN
3010 AA; 327017 MM; P
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Best Local Similarity 77.0
Matches 157; Conservative
                                                                                                                                                                                                                                                                                                                                                            369
1083
11107
11165
1237
1319
196
209
234
250
417
4417
                                                                                                                                                                                                                                                                                                                                                          TRANSMEM
ACT_SITE
ACT_SITE
ACT_SITE
NP_BIND
SITE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                      CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                 CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                              CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CARBOHYD
                                                                                                                                                                                                                                                                                                  CHAIN
                                                                                                                                                                                                                                                                                                                          CHAIN
    a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      염
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ö
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are not restrictions on its use by non-profit institutions as long as its cnotent is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              MEDLINE-92044440; PubMed-1658196;
MISCHORIDA A., MISCHORID M.;
MACHIGA A., MISCHORI M.;
MUCLECTIOR ENGINE:
To Gen. Virol. 72:2697-2704(1991).
To Gen. Virol. 72:26
                                                                                                                                                                                                                          01-AUG-1992 (Rel. 23, Created)
01-AUG-1992 (Rel. 23, Last sequence update)
01-AUG-1992 (Rel. 23, Last sequence update)
01-AUG-1992 (Rel. 41, Last amondation update)
Genome polyprotein [Contains Capsid protein C (Core protein) (P22);
Envelope glycoprotein E1 (GP32) (GP35); Envelope glycoprotein E2
(GP68) (GP70) (NS1); Protein P1 Northuctural protein NS2 (RS2) (Repairtin)
(EC 3.4.21.98); Nonstructural protein NS5A (P6); Nonstructural protein
NS4B (P27); Nonstructural protein NS5A (P56); Nonstructural protein
NS5B (P60) (P70) (RNA-directed RNA pollymerase) (EC 2.7.7.48)].
Hepatitis C virus (isolate HR-J6) (HCV).
Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (RNA)(N).
-I-SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLECCAPSID COVERED BY LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS: PROTEIN M AND GLYCOPROTEIN E. THE NUCLECCAPSID IS A COMPLEX OF PROTEIN C AND MRNA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY S29
                                                                                                                                                                             PRT; 3033 A.A.
1185 CTRGVAKAVDFIPVESMETTMRSP 1208
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Interpro; IPR002166; HCV_RGRP.
Interpro; IPR001650; Hellcase_C.
Interpro; IPR007095; RNA_pol_DS_PS.
Interpro; IPR007094; RNA_pol_PSvir.
Pfam; PF01543; HCV_capsid; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 HCV_NS3.
HCV_NS4a.
HCV_NS4b.
HCV_NS5a.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           EMBL; D00944; BAA00792.1; -.
                                                                                                                                                                             STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       InterPro; IPR001410;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               InterPro; IPR002519;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             InterPro; IPR000745;
InterPro; IPR001490;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            InterPro; IPR002522;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      InterPro; IPR002518;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        InterPro; IPR004109;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             JQ1303; JQ1303.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        THEI.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                NCBI_TaxID-11113;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  HSSP; P27958; IHE
MEROPS; S29.001;
                                                                                                                                                                       POLG_HCVJ6
P26660;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Hepacivirus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  MEROPS
                                                                                                                                                                                셤
```

```
ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      79 IASPKGPVTQMYTNVDKDLVGWQAPQGSRSLTPCTCGSSDL/YLVTRHADVIPVRRRGDSR 138
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 R Pfam; PF01539; HCV_env; 1.

R Pfam; PF01589; HCV_NS1: 1.

R Pfam; PF01589; HCV_NS2: 1.

R Pfam; PF01589; HCV_NS3: 1.

R Pfam; PF01000; HCV_NS4: 1.

R Pfam; PF01000; HCV_NS4: 1.

R Pfam; PF01001; HcV_NS4: 1.

R Pfam; PF01001; HcV_NS5: 1.

R Pfam; PF00271; Helicase_C: 1.

R Pfam; PF00271; Helicase_C: 1.

R Pfam; PF00298; Viral_RRP; 1.

R Probom; PD18602; HCV_NS1: 1.

R Probom; PO17F0101; DF010101; Hydrolase; ATP-Dinding; MTranamembrane; Nonstructural protein; Hydrolase; Serline protease.

TINIT_MET INIT_MET C BY THE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       19 TAYAQQIRGEEGCQETSQIGRDKNQVEGEVQIVSTATQIFLATCINGVCWIVYHGAGIRT 78
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   139 GSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVCTRGVAKAVDFIPVESLETIMRSP 197
                                                                                                                                                                                    CELULAR AMINOPEPTIDASE.
CAPSID PROTEIN (POTEWITAL).
MAJOR ENVELOPE PROTEIN E (POTEWITAL).
NONSTRUCTURAL PROTEIN NI (POTEWITAL).
NONSTRUCTURAL PROTEIN NI (POTEWITAL).
PROTEASE/HELICASE NI (POTEWITAL).
NONSTRUCTURAL PROTEIN NS2 (POTEWITAL).
NONSTRUCTURAL PROTEIN NS48 (POTEWITAL).
NONSTRUCTURAL PROTEIN NS48 (POTEWITAL).
RNA-DIRECTED RNA POLYMERASE (POTEWITAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  0; Gaps
                                                                                                                                                                                                                                                                                                                 CHARGE RELAY SYSTEM (BY SIMILARITY).
CHARGE RELAY SYSTEM (BY SIMILARITY).
CHARGE RELAY SYSTEM (BY SIMILARITY).
                                                                                                                                                                                                                                                                                                                                                                                                                                  (POTENTIAL).
(POTENTIAL).
(POTENTIAL).
(POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                        (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          65.6%; Score 677; DB 1; Length 3033; 68.7%; Pred. No. 3e-56; iive 26; Mismatches 30; Indels (
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   F957F5C1A273BE9E CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                 (GLCNAC. . .)
                                                                                                                                                                                                                                                                                                                                                                                                           (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                             (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                          (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (GLCNAC
                                                                                                                                                                                                                                                                                                                                                    (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                         N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                            N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
                                                                                                                                                                                                                                                                                                        POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                DECH BOX.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    329165 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 123; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            2038
2811
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      649 64
1091 109
2038 203
2811 281
3033 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Query Match
Best Local Similarity
Matches 123; Conserv
                                                                                                                                                                                                                                                                                           2018
                                                                                                                                                                                                                                                                                                                                                                                                                      CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                      TRANSMEM
ACT_SITE
ACT_SITE
ACT_SITE
NP_BIND
SITE
                                                                                                                                                                                                                                                                                                                                                                                     CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                            CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                        CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                            CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                           CARBOHYD
Pfam;
Pfam;
Pfam;
Pfam;
Pfam;
Pfam;
Pfam;
                                                                                                                                                                                                                         CHAIN
                                                                                                                                                                                                                                               CHAIN
                                                                                                                                                                                                   CHAIN
                                                                                                                                                                                                                                                                                            CHAIN
                                                                                                                                                                                                                                     CHAIN
 셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
```

3033 AA

PRT;

STANDARD;

RESULT 8
POLG\_HCVJ8
ID POLG\_HCVJ8
AC P26661;

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-ch).
01-AUG-1992 (Rel. 23, Created)
01-AUG-1992 (Rel. 23, Last sequence update)
28-FEBS-2003 (Rel. 41, Last annotation update)
Genome polyprotein (Contains: Capsid protein C (Core protein) (P22);
Envelope glycoprotein E1 (GP32) (GP35); Envelope glycoprotein E2 (GP68) (GP70) (MS1); Protein P7; Nonstructural protein NS2 (P21)
(EC 3.4.22.-); Protease/helicase NS3 (P70) (Hepacivirin)
(EC 3.4.21.98); Nonstructural protein NS4A (P4); Nonstructural protein NS4B (P27); Nonstructural protein NS5A (P56); Nonstructural protein NS5A (P66) (P70) (RMa-directed RNA polymerase) (EC 2.7.7.48)).
Hepatitis C virus (isolate HG-J8) (HCV).
Viruses; SSRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                                         VITOLOGY 188:331-341(1992).

-!- FUNCTION: THE SMALL PROTEINS NSZA, NSZB, NS4A AND NS4B ARE HYDROPHOBIC, SUGGESTING A POSSIBLE MEMBRANE-RELATED FUNCTION.

NS3 AND NS5 MAY PLAY A ROLE IN THE VIRAL RNA REPLICATION.

-!- CATALYTIC ACTIVITY: Hydrolysis of four peptide bonds in the viral precursor polyprotein, commonly with Asp or Glu in the P6 position, Cys or Thr in P1 and Ser or Ala in P1'.

-!- CATALYTIC ACTIVITY: N nucleoside triphosphate - N diphosphate +
                                                                                                                                                                                                                                                                                                MEDLINE-92230232; PubMed-1314459;
Okamoto H., Kurai K., Okada S.-I., Yamamoto K., Lizuka H., Tanaka T.,
Fukuda S., Tsuda F., Mishiro S.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          -:- SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MINA.
                                                                                                                                                                                                                                                                                                                                                          *Full-length sequence of a hepatitis C virus genome having poor homology to reported isolates: comparative study of four distinct
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   -!- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY S29.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ); IPR007095; RNA_POI_DS_PS.
); IPR007094; RNA_POI_DS_PS.
P01543; HCV_capsid; 1.
P01543; HCV_care; 1.
P01539; HCV_env; 1.
P01539; HCV_NS1; 1.
P01539; HCV_NS2; 1.
P01539; HCV_NS3; 1.
P01539; HCV_NS3; 1.
P01539; HCV_NS3; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              HCV_capsid.
HCV_core.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         HCV_NS4a.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               HCV_NS5a.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             EMBL; D10988; BAA01761.1; -. PIR; A40250; GNWVJ8.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    HCV_NS4b; 1.
HCV_NS5a; 1.
Viral_RdRP; 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            InterPro; IPR001410;
InterPro; IPR002522;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                InterPro; IPR002519;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   InterPro; IPR002531;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    InterPro; IPR002518;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             InterPro; IPR000745;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            InterPro; IPR001490;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       InterPro; IPR004109;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     HSSP; P27958; 1HEI.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           InterPro; IPR00252
                                                                                                                                                                                                                                                                                   SEQUENCE FROM N.A.
                                                                                                                                                                                                                                           WCBI_TaxID=11115;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      MEROPS; S29.001
MEROPS; U39.001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     PF01506;
PF00998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 PF02907;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (RNA)(N).
                                                                                                                                                                                                                             Hepacivirus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                InterPro;
                                                                                                                                                                                                                                                                                                                                                                                            genotypes.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 nterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      interPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          itan;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Pfam;
```

SEQUENCE FROM N.A.
Lensch M.H., Sokolenko A., Herrmann R.G.;
Identification and characterization of the chloroplast HhoA protease,
a homolog to the bacterial periplasmic protease HhoA.";
Submitted (DEC-1998) to the EMBL/GenBank/DDBJ databases.

II; Brassicales; Brassicaceae; Arabidopsis.

NCBI\_TaxID=3702;

```
ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   79 IASPKGPVTQMYINVDKDLVGWQAPQGSRSLTPCTCGSSDLYLVTRHADVIPVRRRGDSR 138
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       78
                                                                                  CASID PROTEIN ( (POTENTIAL).
MATRIX PROTEIN ( POTENTIAL).
MAJOR EWRELOPE PROTEIN ( POTENTIAL).
MAJOR EWRELOPE PROTEIN NSI (POTENTIAL).
NONSTRUCTURAL PROTEIN NSI (POTENTIAL).
PROTEASE/HELICASE NSI (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4A (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4B (POTENTIAL).
RNA-DIRECTED RNA POLYMERASE (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      19 TAYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVSTATQTFLATCINGVCWTVYHGAGTRT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Gaps
     Productions, recovered.

PROMOMERT: SM00487; DEXDC: 1.

Polyprotein: Glycoprotein: Transferase; RNA-directed RNA polymerase; Core protein: Goat protein: Browlope protein: Half-binding; Transmembrane; Nonstructural protein; Hydrolase; Serine protease. INIT_MET 1. REMOVED FROM CAPSID PROTEIN C BY THE CELLULAR AMINOPEPTIONSE.
                                                                                                                                                                                                             CHARGE RELAY SYSTEM (BY SIMILARITY).
CHARGE RELAY SYSTEM (BY SIMILARITY).
CHARGE RELAY SYSTEM (BY SIMILARITY).
                                                                                                                                                                                                                                                                                                             (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                           (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ;
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                    (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        (POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Length 3033;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              31; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                        N-LINKED (GLCNAC...) (POTE N-LINKED (GLCNAC...) (POTE GLCNAC...) (POTE GLCNAC...) (POTE GLCNAC...) (POTE GLCNAC...) (POTE M-LINKED (GLCNAC...) (POTE M-LINKED (GLCNAC...) (POTE M-LINKED (GLCNAC...)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        HHOA_ARATH STANDARD; PRT; 321 AA.
09SEL7; 049507;
16-0CT-2001 (Rel. 40, Created)
16-0CT-2001 (Rel. 40, Last sequence update)
15-SEP-2003 (Rel. 42, Last annotation update)
Protease HhoA, chloroplast precursor (EC 3.4.21.-).
HHOA OR AT4G18370 OR F28312.30.
Arabidopsis thaliana (Mouse-ear cress).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ; DB 1;
4.6e-56;
                                                                                                                                                                                                                                                                      N-LINKED (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                N-LINKED (GLCNAC,
N-LINKED (GLCNAC,
N-LINKED (GLCNAC,
N-LINKED (GLCNAC,
N-LINKED (GLCNAC,
                                                                                                                                                                                                                                                                                                                                                                                                                                                               N-LINKED (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                      (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                   (GLCNAC.
                                                                                                                                                                                                                                         ATP (POTENTIAL)
DECH ROV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  : Pred. No. 4.6e
24; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Score 675;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           2811
a: 330177 MW;
ProDom; PD186062; HCV_NS1; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    65.48;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            124; Conservative
                                                                                                                                                                                                                                                                                                                                     417
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Š
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                               649
1091
2038
2359
2811
3033
                                                                                                                                                                                                                                                                                                                                                                                                  CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                               CARBOHYD
                                                                                                                                                                                                           ACT_SITE
ACT_SITE
ACT_SITE
NP_BIND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CARBOHYD
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                           CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                            CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                  CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CARBOHYD
                                                                                                                                                                                                                                                                         CARBOHYD
                                                                                                                                                                                                                                                                                                  CARBOHYD
                                                                                                                                                                                                                                                                                                               CARBOHYD
                                                                                                                                                                                                                                                                                                                          CARBOHYD
                                                                                                                                                                                                                                                                                                                                      CARBOHYD
                                                                                                                                                                                                                                                                                                                                                   CARBOHYD
                                                                                                                                                                                                                                                                                                                                                              CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                        CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                       CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Query Match
                                                                                                                                                                                                   TRANSMEM
                                                                                                                                                                                                                                                                                      CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Local
                                                                                                                          CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               HHOA ARATH
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Best Loca
Matches
                                                                                                                                         CHAIN
                                                                                                                                                   CHAIN
                                                                                                                                                               CHAIN
                                                                                                                                                                                       CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   RESULT 9
qd
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            SC SE DI LA CO
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
```

```
We like the columnia of the co
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      "The chloroplast lumen from Arabidopsis thallana."; .
Submitted (JuL-201) to the SWISS-PROT data bank.
-1. SUBCELLULAR LOCATION: Chloroplast: within the thylakoid lumen.
-1. SIMILARITY: BELONGS TO PEPTIDASE FAMILY $2C.
-1. CAUTION: Ref.2 sequences differ from that shown due to erroneous gene model prediction. AT4G18370 and AT4G18375 were originally fused into a single gene.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SEQUENCE OF 72-82; 96-110; 150-159; 178-211 AND 306-320. Schubert M., Peterson U., Funk C., Haas B., Schroeder W.P., Kieselbach T.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Nature 402:769-777(1999)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             thaliana.
```

Arabidopsis thaliana (Mouse-ear cress). Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;

```
- I - SIMILARITY: BELONGS TO THE TP013X FAMILY OF LIPOPROTEINS
                                                                                                                                                                                                                                                           485 AA;
                                                                                                                                                                                                                                                                                                 Best Local Similarity
                                                                                                                                                       Complete proteome
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             NCBI_TaxID=9606;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  TISSUE-Brain
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          AAMP_HUMAN
Q13685;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                            152
                                                                                                                                                                                                                                                               SECUENCE
                                                                                                                                                                                                                                                                                     Query Match
                                                                                                                                                                   SIGNAL
                                                                                                                                                                                                     DOMAIN
                                                                                                                                                                                                                DOMAIN
                                                                                                                                                                                                                                       DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     RESULT 11
AAMP_HUMAN
                                                                                                                                                                                                                                                                                                            Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       合
                                                                                                                                                                                                                                                                                                                                                                                                  D.
                                                                                                                                                                                                                                                                                                                                                                                                                             ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                   a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                   ö
                                                                                                                                                                                                                                                                                                                                                                                8
                                                                                                                                                                                                                                                                                                            14;
           modified and this statement is not removed. Usage by and for commercial entitles requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                       95 KTSPSVVYIEAIELPKTSSGDIL----TDEENGKIEGTGSGFVWDKLGHIVTNYHVIAKL 150
                                                                                                                                                                                                                                                                                                                                                                                54 ATQTF-LATCINGVCWTVYHGAGTRTIASPKGPVTQMYTNVDKDLVGWQAPQGSRSLTPC 112
                                                                                                                                                                                                                                                                                                                                                                                            151
                                                                                                                                                                                                                                                                                                                                                                                                                                                    203 VLGTSNDLRVGQSCFAI----GNPYGYENTLTIGVVSGLGREIPSPNGKSISEAIQTDA 257
                                                                                                                                                                                                                                                                                                                                  2 KKKGSVVIVGRINL----SGDTAYAQQTRGEEGCQETSQTG--RDK--NQVEGEVQIVST 53
                                                                                                                                                                                                                                                                                                                                                                                                                             ---SLLSPRPISYLK----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Fraser C.M., Norris S.J., Weinstock G.M., White O., Sutton G.G., Dodson R., Garin M., Hickey E.K., Clayton R., Ketchum K.A., Khalon B., Hardham J.M., McLeod M.P., Salzberg S., Peterson J., Khalak H., Richardson D., Howell J.K., Chidambaram M., Utterback T., McDonald L., Artiach P., Bowman C., Cotton M.D., Fulii C., Garland S., Hatch B., Horst K., Roberts K., Sandusky M., Weidman J., Smith H.O., Venter J.C.,
                                                                                                                                                                                                                                                                                                             Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Science 281:375-388(1998).
                                                                                                                                                       Serine protease; Chloroplast; Thylakoid; Transit peptide.
1 26 CHLOROPLAST (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Spirochaetes; Spirochaetales; Spirochaetaceae; Treponema
 is in
                                                                                                                                                                                                                                                                                                                                                                                                                                                                            152 ----GSSGGPLLCPAGHAVGIFRAAVCTR---GVAKAVDF-IPVESLETTM 194
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        258 DINSGNSGGPLLDSYGHTIGV-NTATFTRKGSGMSSGVNFAIPLDTVVRTV 307
                                                                                                                                                                                                                                                                                                             26;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Complete genome sequence of Treponema pallidum, the syphilis
                                                                                                                                                                                                             CHARGE RELAY SYSTEM (POTENTIAL)
CHARGE RELAY SYSTEM (POTENTIAL)
CHARGE RELAY SYSTEM (POTENTIAL)
                                                                                                                                                                                                                                                                                    Score 86; DB 1; Length 321;
Pred. No. 0.67;
 non-profit institutions as long as its content
                                                                                                                                                                                                                                                                                                          85; Indels
                                                                                                                                                                                                                                                    -> G (IN REF. 1).
68DB81E0BD27A7A7 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         136_TREPA STANDARD; PRI; 485 AA. 083172; 16-0CT-2001 (Rel. 40, Created) 16-0CT-2001 (Rel. 40, Last sequence update) 16-0CT-2001 (Rel. 40, Last annotation update) Hypothetical lipoprotein TP0136 precursor.
                                                                                                                                                                                       PROTEASE HHOA. POLY-GLU.
                                                                                                                                                                                                                                                                                                            28; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                              113 TCGSSDLYLVTRHADVIPVRRRGDSRG----
                                                                                                                                                                              THYLAKOID
                                                                                                      InterPro; IPR001940; Proteasc2C.
InterPro; IPR001284; Scr_protease_Try.
Pfam; PF000089; trypsin; PFNTES; PR00884; PROTEASESCC.
                                                                       EMBL; AL021710; CAA16717.1; ALT_SEQ
EMBL; AL161548; CAB78839.1; ALT_SEQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       STRAIN-Nichols;
MEDLINE-98332770; PubMed-9665876;
                                                            EMBL; AF114386; AAF24060.1; -
                                                                                                                                                                                                                                                               34691 MW;
                                                                                                                                                                                                                                                                                   8.3%;
                                                                                                                                                                                                                                                                                    Query Match
Best Local Similarity 26.89
Matches 62; Conservative
                                                                                                                                                                             71
321
87
145
186
264
                                                                                                                                                                                                                                                   40
321 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Treponema pallidum
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SEQUENCE FROM N.A.
                                                                                              MEROPS; S01.279;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  NCBI_TaxID=160;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            (Potential)
                                                                                                                                                         Hydrolase;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Bacteria;
                                                                                                                                                                                                                                       ACT_SITE
CONFLICT
                                                                                                                                                                                                                 ACT_SITE
                                                                                                                                                                                                                              ACT_SITE
                                                                                                                                                                                                                                                               SECTENCE
                                                                                                                                                                             TRANSIT
                                                                                                                                                                   PRANSIT
                                                                                                                                                                                           CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                (136_TREPA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ð
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entitles requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@lsb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  54 AGSKLYATNGRLWEKELNGTGSWQKVSSSSVPTDSDK----KVMSIATDGNTFVLACVP 108
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           16 SGDTAYA-----QQTRGEEGCQETSQ----TGRDKNQVEGEVQIVSTATQTFLATCI- 63
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    64 -NGVCWTVYHGAG---TRTIASPKGPVTQMYTNVDKDLVG-----WQAPQGSRSLTPCT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         109 GTGVYKHCVNGAGSSSTGTTASPSTETCSQHAT ---- LVGGTSKPFWLVPGGTGNNGNCG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          114 C----GSSDLYLVTRHADVIP-----VRRRGDSRGSLLSPRPISYLK-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Eukaryota, Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                71;
                                                                                                                                                                                                                                                                                                                                              HYPOTHETICAL LIPOPROTEIN TP0136.
N-ACYL DIGLYCERIDE (POTENTIAL).
GLY/SER-RICH.
GLZ/SER-RICH.
GLZ/SER-RICH.
POLY-SER.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Length 485;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              48984 MW; C7A4CEEDC7DC5CED CRC64;
                                                                                                                                                                                                                                                                  Hypothetical protein; Lipoprotein; Membrane; Signal;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   16;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Score 84.5; DB 1;
Pred. No. 1.5;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   225 DMGEGYVVTTKHLYTKNGSSSAGPAQCPGGGGGG 258
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  -----GSSGGPLLCPAGHAVG 167
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 15-JUL-1998 (Rel. 36, Created)
15-JUL-1998 (Rel. 36, Last sequence update)
28-FEB-2003 (Rel. 41, Last annotation update)
Anglo-associated migratory cell protein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  452 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                17; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SEQUENCE FROM N.A., AND PARTIAL SEQUENCE
                                                                                                                                                                                                                                                                                                                     POTENTIAL
                                                                                                                                                                                                              EMBL; AE001199; AAC65137.1; ALT_INIT.
TIGR; TP0136; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      POLY-SER
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    MEDLINE-95262124; PubMed-7743515;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   50; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Homo sapiens (Human)
```

MEDLINE-20363099; PubMed-10907853;

```
between the Swiss Institute of Bloinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          6
 This SWISS-PROT entry is copyright. It is produced through a collaboration
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        240
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          WIVYHGAGIRIIASPKGPVTQMYINVDKDLVGWQAPQGSRSL----TPCTCGSSDLYLV 122
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      123 TRHADVIPVRRR---GDSRGS----LLSPRPISYLKGSSG--GPLLCPA----- 162
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    241 ----VLPDGKRAVVGYEDGTIRIWDLRQGSPIHVLKGTEGHQGPLTCVAANQDGSLILT 295
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         197 WMEWH----PRAPVLLAGT-ADGNTWMWKVPNGDCKTFQGPNCPATCGR----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
eurosids II; Brassicales; Brassicaceae; Arabidopsis.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SEQUENCE FROM N.A., AND CHARACTERIZATION.
MEDLINE-98175982; PubMed-9507020;
Itzhaki H., Naveh L., Lindahl M., Cook M., Adam Z.;
Identification and characterization of Degp, a serine protease associated with the luminal side of the thylakoid membrane.";
J. Biol. Chem. 273:7094-7098(1998).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     163 -----GHAVGIFR----AAVCTRGVAKAVDFIPVESL 190
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                296 GSVDCQAKLVSATTGKVVGVFRPETVASQPSLGEGEESESNSVESL 341
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ; Pred. No. 3;
13; Mismatches 47; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        7.8%; Score 81; DB 1; Length 452; 25.3%; Pred. No. 3;
                                                                                                                                                                                                                                                                                                                     HEPARIN-BINDING (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             DEGI_ARATH STANDARD; PRT; 437 AA.
022609, 09LKB5;
16-0CT-2001 (Rel. 40, Last sequence update)
16-0CT-2003 (Rel. 40, Last annotation update)
28-FED-2003 (Rel. 41, Last annotation update)
Protease Do-like 1, chloroplast precursor (EC 3.4.21.-).
DEGPI_OR DEGP_OR AT3G27925 OR K16N12.18.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                            DA1413D25EB236C0 CRC64;
                                                                                                                                                                    MIN, 603488; ...

MIN, 603488; ...

GO; GO:0008201; F:heparin binding activity; TAS.

InterPro; PRO0160; WD40; WD40.

SMART; SM00320; WD40; 8.

SMART; SM00320; WD40; 8.

PROSITE; PS00082; WD_REPEATS_1; 1.

PROSITE; PS50082; WD_REPEATS_2; 6.

PROSITE; PS50294; WD_REPEATS_REGION; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Arabidopsis thaliana (Mouse-ear cress).
                                                                                                                                                                                                                                                                                                                                 POLY-GLU.
WD 1.
WD 3.
WD 4.
WD 5.
WD 6.
WD 6.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                             49015 MW;
                                                                                                                          EMBL; M95627; AAA68889.1; -. PIR; I39383; I39383.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Local Similarity 25.39 tes 42; Conservative
                                                                                                                                                                                                                                                                                                                   18
77
138
180
220
                                                                                                                                                        HGNC:18; AAMP.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       STRAIN-CV. Columbia;
                                                                                                                                                                                                                                                                                                                                                                                               231
276
333
374
416
452 AA;
                                                                                                                                                                                                                                                                                                    Repeat; WD repeat.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SEQUENCE FROM N.A
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        NCBI_TaxID-3702;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          89
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SEQUENCE
                                                                                                                                                                                                                                                                                                                                   DOMAIN
REPEAT
REPEAT
REPEAT
                                                                                                                                                             Genew;
                                                                                                                                                                                                                                                                                                                                                                                               REPEAT
REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                               REPEAT
REPEAT
                                                                                                                                                                                                                                                                                                                     DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Best Loc
Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
```

```
..
                                                                                                                                                                                                                                                                                                    This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bloinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            102
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   150 VPQGSGSGFVWDKQGHIVTNYHVIRGASDLRVTLADQTTFDAKVVGFDQDKDVAVLRIDA 209
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    -- TRHADVIPVRRRGDSRGSLLSPRPI 147
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          70 VYHGAGTRIIASPKGPVIQMY-----QA
                                                                                                                                                                                            ENZYME REGULATION: INHIBITED BY PHENYLMETHYLSULFONYL FLUORIDE AND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Gaps
                                                                                                          STRAIN-cv. Columbia;
Kieselbach T., Bystedt M., Schroeder W.P.;
Submitted (JUL-2000) to the SWISS-PROT data bank.
-!- FUNCTION: SERINE PROTEASE THAT IS REQUIRED AT HIGH TEMPERATURE.
MAY BE INVOLVED IN THE DEGRADATION OF DAMAGED PROTEINS. IN VIVO,
CAN DEGRADE BETA-CASEIN.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CHARGE RELAY SYSTEM (POTENTIAL).
CHARGE RELAY SYSTEM (POTENTIAL).
CHARGE RELAY SYSTEM (POTENTIAL).
HSPPSSQLSNST -> SSTLFLHSPPSSHL (IN REF.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Transit peptide; Chloroplast; Thylakoid.
           Kaneko T., Katoh T., Sato S., Nakamura A., Asamizu E:, Tabata S.;
*Structural analysis of Arabidopsis thaliana chromosome 3. II.
Sequence features of the 4,251,695 bp regions covered by 90 P1, TAC
                                                                                                                                                                                                                        -!- SUBCELLULAR LOCATION: BOUND TO LUMINAL SIDE OF THE THYLAKOID
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                55;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  DB 1; Length 437;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             56; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1497BlAB3F5FF2A4 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CHLOROPLAST (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       V -> 1 (IN REF. 2).
P -> S (IN REF. 2).
G -> R (IN REF. 2).
L -> HF (IN REF. 2).
L -> HF (IN REF. 2).
L -> V (IN REF. 2).
I -> V (IN REF. 2).
O -> E (IN REF. 2).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              PROTEASE DO-LIKE 1.
                                                                                                                                                                                                                                                              -:- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S2C.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SERINE PROTEASE.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  7.6%; Score 78.5;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 .
No.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 THYLAKOID
                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Interpro: IPR001478; PD2.
Interpro: IPR001440; Protease2C.
InterPro: IPR001244; Ser_protease_Try.
Pfam: PF00089; PD2; 1.
Pfam: PF00089; Irypsin; 1.
PRINTS: PR00834; PROTEASES2C.
SMART; SM00228; PD2; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                EMBL; AF028842; AAC39436.1; -.
EMBL; AP000371; BAB02539.1; -.
EMBL; AP001302; BAB02539.1; JOINED.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Pred.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  103 PQGSRSLTPCTCGSSDLYLV----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              17;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        46213 MW;
                                                                                                                                                                                                                                                 INDUCTION: By heat shock
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Hydrolase; Serine protease;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Query Match
Best Local Similarity 25.6%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        PROSITE; PS50106; PDZ; 1.
                                                                   Res. 7:217-221(2000).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          36
64
64
64
64
83
83
81
81
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          321
421
171
2201
230
23
                                                                                                                                                                                                           O-PHENANTHROLINE
                                                                                              SEQUENCE OF 104-118.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         416
437 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                           MEROPS; S01.279;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              104
152
324
171
201
280
                                                      BAC clones.'
                                                                                                                                                                                                                                       MEMBRANE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                44;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ACT_SITE
ACT_SITE
CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CONFLICT
CONFLICT
CONFLICT
CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ACT_SITE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 RANSIT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    FRANSIT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CHAIN
DOMAIN
DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Matches
Š
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     집
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
```

```
CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                     CARBOHYD
                                                                                                                                                                      ACT_SITE
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                        Query Match
                                                                                                                                                                                                 CARBOHYD
                                                                                                                                                                                                            CARBOHYD
                                                                                                                                                                                                                         CARBOHYD
                                                                                                                                                                                                                                     CARBOHYD
                                                                                                                                                                                                                                                   CARBOHYD
                                                                                                                                                                                                                                                               CARBOHYD
                                                                                                                                                                                                                                                                           CARBOHYD
                                                                                                                                                                                                                                                                                       CARBOHYD
                                                                                                                                                                                                                                                                                                   CARBOHYD
                                                                                                                                                                                                                                                                                                                 CARBOHYD
                                                                                                                                                                                                                                                                                                                             CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                 CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                              CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                          CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                SEQUENCE
                                                                                                                                                 DOMAIN
                                                                                                          DOMAIN
                                                       DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CTRL_HUMAN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entitles requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
210 PK--NKLRPIPVGVSADLLVGQKVFAIGNPFGLDHTLTTGVISGLRREIS--SAATGRPI 265
                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CATALYTIC ACTIVITY: Protein tyrosine phosphate + H(2)0 = protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SÜBCELLULAR LOCATION: Type I membrane protein.
DEVELCOMEKNAL STRAE: DÄTECTABLE IN THE BIBLAST OF COCYTES AND
THROGGHOUT EARLY MOUSE EMBRYO DEVELOPMENT. IN ADULY, EXPRESSION
LOCALIZED IN GONADAL GERM CELLS.
SIMILARITY: Contains 2 protein tyrosine phosphatase domains.
SIMILARITY: Contains 10 fibronectin type III domains.
                                                                                                                                                                                                                             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalla; Eutheria; Rodentia; Sciurognathi; Muridae; Musinae; Mus
NCBL_TaxID=10090;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 EMBRYONIC STEM CELL PROTEIN TYROSINE
                            SYL-----KGSSGGPLLCPAGHAVGIFRAAVCTRGVAKAVDF-IPVESL 190
                                                                                                                                                                                                                                                                                                                                                                                                                Lee K., Nichols J., Smith A.;
wech. Dev. 61:131-215(1996).
-!- FUNCTION: MAY PLAY A ROLE IN THE MAINTENANCE OF PLURIPOTENCY.
DOWN-REGULATED DURING DIFFERENTIATION.
                                                                                                                                                                                                                                                                                               TISSUE-Embryonic stem cells;
WEDLINE-97109513; PubMed-805.1793;
Lee K., Nichols J., Smith A.;
"Identification of a developmentally regulated protein tyrosine phosphatase in embryonic stem cells that is a marker of pluripotential epiblast and early mesoderm.";
Wech. Dev. 59:153-164(1996).
                                         266 ODVIQTDAAINPGNSGGPLLDSSGTLIGINTAIYSPSGASSGVGFSIPVDTV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Pfam; PF00041; fn3; 7.

Pfam; PF00041; fn3; 7.

Pfam; PF00102; v_phosphatase; 1.

R PRINTS; PR00700; PRTYPHPHTASE.

SMART; SM00060; FN3; 8.

SMART; SM000194; PTPC; 1.

DR PROSTTE; PS00194; TYR_PHOSPHATASE_1; 1.

DR PROSTTE; PS50056; TYR_PHOSPHATASE_2; 1.

DR PROSTTE; PS50055; TYR_PHOSPHATASE_2; 1.

W Hydrolase; Transmembrane; Repeat; Signal; Glycoprotein.
                                                                                                                                          15-JUL-1998 (Rel. 36, Created)
15-JUL-1998 (Rel. 36, Last sequence update)
16-JCT-2001 (Rel. 40, Last annotation update)
Embryonic stem cell protein tyrosine phosphatase precursor (EC 3.1.3.48) (ES cell phosphatase).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         EXTRACELLULAR (POTENTIAL)
POTENTIAL.
                                                                                                                    PRT; 1705 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                PHOSPHATASE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             MGD; MGI:108027; ESP.
InterPro; IPR003961; FN_III.
InterPro; IPR000387; TYR_phosphatase.
InterPro; IPR000242; TYR_PP.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     EMBL; U36488; AAC52868.1; -. HSSP; P18052; 1YFO.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                               tyrosine + phosphate.
                                                                                                                    STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1077
                                                                                                                                                                                                                       musculus (Mouse)
                                                                                                                                                                                                                                                                                      SEQUENCE FROM N.A
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           19
1078
                                                                                                                    PTPO_MOUSE
P70289;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         DOMAIN
TRANSMEM
                            148
                                                                                                                                                                                                                                                                                                                                                                                                       ERRATUM.
                                                                                                        PTPO_MOUSE
                                                                                           RESULT 13
                                                                                                                                  ö
```

```
е
::
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               -----NVDKDL-VGW 100
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   442 TFCGLVPGAHYRV-----DIASSMGDITQSLTGYTSPLPPQSLEIISRNSPSDLTIGW 494
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          9 IVGRINLSGDTAYAQQTRGE----EGCQETSQTGR------DKNQVEGEVQIVSTATQT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalla; Eutheria; Primates; Catarrhini; Hominidae; Homo.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Larsen F., Solheim J., Kristensen T., Kolsto A.B., Prydz H.; "A tight cluster of five unrelated human genes on chromosome 16q22.1.";
FIBRONECTIN TYPE-III 1.
FIBRONECTIN TYPE-III 2.
FIBRONECTIN TYPE-III 3.
FIBRONECTIN TYPE-III 4.
FIBRONECTIN TYPE-III 6.
FIBRONECTIN TYPE-III 6.
FIBRONECTIN TYPE-III 6.
FIBRONECTIN TYPE-III 7.
FIBRONECTIN TYPE-III 9.
FIBRONECTIN TYPE-III 9.
FIBRONECTIN TYPE-III 9.
FIBRONECTIN TYPE-III 1 9.
FUBRONECTIN TYPE-III 9.
FUBRONECTIN TY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             7.6%; Score 78; DB 1; Length 1705; 55%; Pred. No. 28; Ve 12; Mismatches 53; Indels 9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              101 -QAPQGSRSLTPCTCGSSDLYLVTRHADVIPVRRRGDSRGSLLSP 144
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             495 APAP------GOMEGYKVTWHQD-------GSORSP 517
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Chymotrypsin-like protease CTRL-1 precursor (EC 3.4.21.-). CTRL OR CTRL1. Homo sapiens (Human).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  2783755F15387D5B CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            (GICNAC. . .)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     GLCNAC. . .)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              (GLCNAC. . .)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          (GLCNAC. . .)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    GLCNAC. ..)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (GLCNAC. . . )
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            (GLCNAC. . . .
                                                                                                                                                                                                                                                                                                                                                                             BY SIMILARITY.
N-LINKED (GLCNAC.
N-LINKED (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Hum. Mol. Genet. 2:1589-1595(1993).
-!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            01-FEB-1995 (Rel. 31, Created)
01-FEB-1995 (Rel. 31, Last sequence update)
28-FEB-2003 (Rel. 41, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                             GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    N-LINKED (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    58 FLATCINGVCWTVYHGAGTRTIASPKGPVTQMYT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                      N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SEQUENCE FROM N.A. MEDLINE-94093544; PubMed-8268911;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     186795 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ilarity 25.5%;
Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    STANDARD;
                                                                                                                                                                                                                                                                                    905
11418
1700
1350
1350
1117
1117
1239
2239
2299
3431
570
620
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      882 88
970 97
982 98
1705 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Best Local Similarity
Matches 42; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          NCBI_TaxID-9606;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CTRL_HUMAN
P40313;
```

MEDLINE-89073750; PubMed-3201749;

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation-the Buropean Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                         ACTIVATION.

ACTIVATION
CHYNOTRYPSIN-LIKE PROTEASE CTRL-1.
CHARGE RELAY SYSTEM (BY SIMILARITY).
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       DB 1; Length 264;
                                                                                                                                                                                                                                                                                                         InterPro; IPRU0123*, ....r.
Pfam: PROM089; trypsin: 1.

PRINTS; PR00722; CHYMCTMFEIN.

SMART; SM00020; Tryp_SPC; 1.

PR05TE; PS50240; TRYPSIN_DOM; 1.

PR05TE; PS00135; TRYPSIN_HIS; 1.

PR05TE; PS00135; TRYPSIN_ESR; 1.

Hydrolase; Serine protease; Glycoprotein; Zymogen; Signal.
                                                                                                                               EMBL; X71874; CAA50710.1; -.
EMBL; X71877; CAA50711.1; -.
PIR; 138136; 138136.
HSPS; P00753; LDP0.
MEROPS; S01.256; -.
Genew; HGNC:2524; CTRL.
MIN; 118888; -.
GO; GO:0005565; C:extracellular space; TAS.
GO; GO:0005686; P:dfgestion; TAS.
GO; GO:0005688; P:proteolysis and peptidolysis; TAS.
InterPro; IPR001314; Chymotrypsin.
InterPro; IPR001254; Ser_protease_Try.
Pfam; PF00089; trysin; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SIMILARITY, 3F629F02FA6DDFB4 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      28002 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       210
264 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                               PROPEP
CHAIN
ACT_SITE
ACT_SITE
ACT_SITE
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         DISULFID
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           DISULFID
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Query Match
Best Local 8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        DISULFID
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            DISULFID
```

```
44 VEGEVQIVSTATQTFLATCINGVCWTVYHGAGTRTIASPKGPVTQMYTNVDKDLVGWQAP 103
                                                                                                      118 MNNDVTLLKLASPAQYTTRISPVC------LASSNEALTEGLTCV---TTGWGRL 163
                                                                                                                                                                  104 QGSRSLTPCTCGSSDLYLVTRHADVIPVRRRGDSRGSLLSPRPI----SYLKGSSGG
                                    51; Indels 31; Gaps
                  4.3;
                                    18; Mismatches
7.4%; Score 76.5;
25.9%; Pred. No. 4.
                                                                                                                                                                                                                157 PLLCPAGHA --- VGI 168
                                                                                                                                                                                                                                         217 PLVCQKGNTWVLIGI 231
                                    Conservative
                 Similarity
                                    Matches
                                                                                                        g
                                                                                                                                            ò
                                                                                                                                                                                                                ö
```

```
Squirrel monkey retrovirus (SMRV'H) (SMRV-HLB).
Viruses; Retroid viruses; Retroviridae; Betaretrovirus.
                   01-MAY 1991 (Rel. 18, Created)
01-MAY-1991 (Rel. 18, Last sequence update)
28-FEB-2003 (Rel. 41, Last annotation update)
PRT.
 STANDARD;
                                                                                                      NCBI_TaxID=11856;
VPRT_SMRVH
P21407;
```

SEQUENCE FROM N.A.

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the SWISS Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            111 NDFEGEIHIILSTTKDL -------VTIPKGTRLAQIVILPLQQINSN 150
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           93 VDKDLVGWQAPQGSRSLTPCTCGSSDLYLV---TRHADVIPVRRRGDSRGSLL----SPR 145
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 85
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         151 FHKPYRGASAP-----GSSDVYWVQQISQQRPTLKLKLNGKLFSGILDTGADAT 199
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               42 NOVEGEVOIVSTATQTFLATCINGVCWTVYHGAGIRTIASPKG-------PVTQMYTN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               47; Gaps
                                                 gene structure
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Length 323;
                                        Oda T., Ikeda S., Watanabe S., Hatsushika M., Akiyama K.,
Mitsunobu F.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 49; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 BY SIMILARITY.
5D6CEA38BA932786 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               DB 1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                5.5;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               16; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               7.4%; Score 76.5;
23.3%; Pred. No. 5.
                                                                                                                                                                                                                                                                                                                                              InterPro: IPR001995; Aspprotease_rtrv.
InterPro: IPR001969; Aspprotease_site.
InterPro: IPR001428; DeoxyUTPase.
InterPro: IPR001428; DeoxyUTPase.
InterPro: IPR001428; G_patch.
Pfam: PF00582; dUTPase; 1.
Pfam: PF00585; G_patch; 1.
Pfam: PF00777; rvp: 1.
PR017F; SM00443; G_patch: 1.
PR051TE; PS00174; ASP_PROTEASE; 1.
PR05ITE; PS50174; G_PATCH: 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           146 PISYLKGSSGGPLLCPAGHAVGIFRA 171
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        200 VISYTHWPRNWPLITVATHLRGIGGA 225
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Search completed: August 30, 2003, 19:13:46 Job time : 10.7567 secs
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 G-PATCH
                                                                                                                                                                                                                                                                                      EMBL; M23385; AAA66452.1; ALT_INIT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              323 AA; 35126 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    protease.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Best Local Similarity 23.3%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               34; Conservative
                                                                                                                                                                                                                                                                                                     PIR; B31827; PRLJHD.
HSSP; P06968; 1EUW.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Hydrolase; Aspartyl
                                                                                                                                                                                                                                                                                                                                      MEROPS; A02.UPW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ACT_SITE
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      q
ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
```

hepatitis hepatitis hepatitis hepatitis hepatitis hepatitis hepatitis

hepatitis hepatitis hepatitis

hepatitis hepatitis hepatitis hepatitis hepatitis hepatitis hepatitis hepatitis hepatitis

091121 091149 091141 091170 091185 091183 091183 091184 091184 091183 091183 091183 091184 091185 091187 091181 091181 091181 091181 091181 091181 091181 091181 091181 091181 091181 091181 091181 091181 091181 091181

hepatitis hepatitis hepatitis hepatitis hepatitis hepatitis

hepatitis

```
MEDLINE-20323484; Pubmed-10864644;
Lai V.C., Zhong W., Skelton A., Ingravallo P., Vassilev V.,
Lai V.C., Long W., Skelton A., Ingravallo P., Vassilev V.,
Donis R.O., Hong Y., Lau J.Y.;
"Generation and characterization of a hepatitis C virus NS3 protease-dependent bovine viral diarrhea virus.";
J. Virol. 74:6339-6347(2000).
SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Lai V.C.H., Hong 2.;
Submitted (MAY-2000) to the EMBL/GenBank/DDBJ databases.
EMBL: #F268278; AAF82566.1; -.
HSSP; P26663; 1JXP.
                                                                                                                                                                                                                                                                                                                                                                                                                                                    Last sequence update)
                                                                                                                                                                                                                                                                                                                                                                                                                   PRT; 4040 AA
                                                                                                                                                                                                                                                                                                                                                            ALIGNMENTS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Interpro: IPR004109; HCV_NS3.
Interpro: IPR002166; HCV_RdRP.
Interpro: IPR001056; Helicase_C.
Interpro: IPR001005; Myb_DNA_binding.
Interpro: IPR001005; RNase_T2.
Interpro: IPR007095; RNA_pol_DS_PS.
Interpro: IPR007094; RNA_pol_DS_PS.
Interpro: IPR007094; RNA_pol_PSvir.
Pfam; PF02907; HCV_NS3; 1.
Pfam; PF00998; Viral_RdRP; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        InterPro; IPR000280; CDvir_endptseP80
InterPro; IPR001410; DEAD.
091RS1
091RR08
091RR10
091RR10
091RR86
091RR86
091RR36
091RR37
091RR37
091RR36
                                                                                                                                                                                                                                                                                                                                                                                                                                          Created)
                                                                                                                                                                                                                                                                                      070818
091RR9
09DTE2
                                                                                                                                                                                                                                                                 070817
                                                                                                                                                                                                                                                                             Q91RT8
                                                                                                                                                                                                                                                                                                                                                                                                                                        01-OCT-2000 (TrEMBLrel. 15, 01-OCT-2000 (TrEMBLrel. 15, 01-MAR-2003 (TrEMBLrel. 23,
 PRELIMINARY;
Genome polyprotein.
Mucosal disease virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Pestivirus.
NCBI_TaxID=11099;
 MEROPS; S31.001
 872
871
870.5
870.8
869.5
869
869
869
869
868.5
868
                                                                                                                                                                                                                                                                                                866.5
866.5
866.5
Q91FH8
                                                                                                                                                                                                                                                                                                                                                                                           RESULT 1
 D9IFH8
                                                                                                                                                                                                                                                                                                                                                                                                                   DORRER REPARENCE OF THE PROPERTY OF THE PROPER
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       091fh8 mucosal dis
036579 hepatitis c
081756 hepatitis c
09165 hepatitis c
09168 hepatitis c
036608 hepatitis c
099ws hepatitis c
091rr8 hepatitis c
091rr8 hepatitis c
091rr5 hepatitis c
091rr5 hepatitis c
091rr5 hepatitis c
091rr7 hepatitis c
091rr7 hepatitis c
091rr7 hepatitis c
                                                                                          (without alignments)
1352.314 Million cell updates/sec
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Pred. No. is the number of results prodicted by chance to have a score greater than or equal to the score of the result being printed,
                                                                                                                                                1 MKKKGSVVIVGRINLSGDTA.....VAKAVDFIPVESLETTMRSP 197
                                                                              ; Search time 37.5921 Seconds
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Description
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  score greater than or equal to the score of the result bein and is derived by analysis of the total score distribution.
                                                                                                                                                                                                                               830525
          GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd
                                                                                                                                                                                                                             Total number of hits satisfying chosen parameters:
                                                                                                                                                                                                        830525 seqs, 258052604 residues
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SUMMARIES
                                                                                                                                                                                                                                                                                                            summaries
                                                        OM protein - protein search, using sw model
                                                                            August 30, 2003, 19:00:22
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Q91RT9
Q91RR3
Q91RR4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          091FH8
036579
081756
091FE5
09ELS8
003463
036608
09PWX5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1RR5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1RR2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1RT5
                                                                                                                                                                        BLOSUM62
Gapop 10.0 , Gapext 0.5
                                                                                                                                                                                                                                                                                                                                                                              sp_human:*
sp_invertebrate:*
sp_marmal:*
sp_mhc:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                sp_unclassified:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     sp_vertebrate:*
                                                                                                                                                                                                                                                                                   Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 st
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          sp_rvirus:*
sp_bacteriap:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      7
                                                                                                                                                                                                                                                                                                                                                                                                                           sp_organelle:*
sp_phage:*
                                                                                                                                                                                                                                                                                                                                              sp_archea:*
sp_bacteria:*
                                                                                                                                                                                                                                                    Minimum DB seq length: 0
Maximum DB seq length: 200000000
                                                                                                                                                                                                                                                                                                                                                                                                                                                 sp_plant:*
sp_rodent:*
sp_virus:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  sp_archeap:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             US-09-965-594-16
1032
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      DB
                                                                                                                                                                                                                                                                                                                                     SPTREMBL_23:*
                                                                                                                                                                                                                                                                                                                                                                      sp_fungi:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          4040
3011
2436
3011
3011
3011
3015
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  181
181
181
181
181
181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Query
Match 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Title:
Perfect score:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Scoring table:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Score
                                                                                                                                                  Sequence:
                                                                                                                                                                                                         Searched:
                                                                                                                                                                                                                                                                                                                                     Database
                                                                                on:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ş
```

```
NCBI_TaxID-11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Q81756
Q81756;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 RESULT 3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                윱
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                 CINGVCWTVYHGAGTRTIASPKGPVTQMYTNVDKDLVGWQAPQGSRSLTPCTCGSSDLYL 121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                          122 VTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVCTRGVAKA 181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              61
                                                                                                                                                                                                                                                                                                                                                  10 GSVVIVGRIVLSGSGSITACAQQTRGLLGCKITSLTGRDKNQVEGEVQIVSTATQTFLAT 69
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Kolykhalov A.A., Agapov E.V., Blight K.J., Mihalik K., Feinstone S.M., Rice C.M.; "Transmission of hepatitis C by intrahepatic inoculation with
                                                                                                                                                                                                                                                                                                           5 GSVVIVGRINLSGD---TAYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVSTATQTFLAT
               SMART: SM00487; DEXDC; 1.
SMART: SM00480; HELICC; 1.
PROSITE; PS00037; WTB=1; 1.
PROSITE; PS50507; RDRP_VOTAL; 1.
PROSITE; PS50521; RDRP_VOTAL; 1.
PROSITE; PS50551; RNRSE_T2_2; 1.
ATP-binding; Helicase; Hydrolase; Nonstructural protein; Polyprotein; RNA-directed RNA polymerase; Transferase.
SEQUENCE 4040 AA; 453073 MW; ADE87791D055B9DC CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Genome polyprotein.
Hepatitis C virus.
Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                   3;
                                                                                                                                                                                                                          87.5%; Score 902.5; DB 12; Length 4040; 90.8%; Pred. No. 3e-81; Live 5; Mismatches 10; Indels 3;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        01-JAN-1998 (TrEMBLrel. 05, Created)
01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   PRT; 3011 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           IPR007095; RNA_pol_DS_PS. IPR007094; RNA_pol_PSvir.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   STRAIN-H77;
MEDLINE-97373636; PubMed-9228008;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        HCV_capsid.
HCV_core.
HCV_env.
HCV_NS1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   HCV_NS3.
HCV_NS4a.
HCV_NS4b.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            HCV_NS5a.
HCV_RdRP.
Helicase_C
PRINTS; PR00729; CDVENDOPTASE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         182 VDFIPVESLETTMRS 196
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               InterPro; IPR001410; DEAD.
InterPro; IPR002522; HCV_C;
InterPro; IPR002521; HCV_C
                                                                                                                                                                                                                                                Best Local Similarity 90.8
Matches 177; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          InterPro; IPR002531;
InterPro; IPR002518;
InterPro; IPR004109;
InterPro; IPR000745;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           IPR001490;
IPR002868;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      IPR002519;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    IPR002166;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        IPR001650;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          NCBI_TaxID-11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Hepacivirus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           InterPro;
InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          InterPro;
                                                                                                                                                                                                                                                                                                                                                                                            62
                                                                                                                                                                                                                              Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                036579
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           RESULT 2
036579
  DR DR DR KW KW COR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       DER PRESENTATION OF THE PRESENT OF THE PROPERTY OF THE PROPERT
                                                                                                                                                                                                                                                                                                             ö
                                                                                                                                                                                                                                                                                                                                                  a
                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                          õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              윱
```

```
PFam; PF01543; HCV_cape...,

DR Pfam; PF01542; HCV_core; 1.

DR Pfam; PF01542; HCV_core; 1.

DR Pfam; PF01539; HCV_env; 1.

DR Pfam; PF01560; HCV_NS2; 1.

DR Pfam; PF01006; HCV_NS4; 1.

DR Pfam; PF01001; HCV_NS4b; 1.

DR Pfam; PF00271; helicase_C: 1.

DR Pfam; PF00271; helicase_C: 1.

DR Pfam; PF00571; helicase_C: 1.

DR Pfam; PF00571; DEXDC: 1.

DR PROSITE; PS50507; RDRP_POSITIVE; 1.

DR PROSITE; PS505
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1125 CGSSDLYLVTRHADVIPVRRRGDGRGSLLSPRPISYLKGSSGGPLLCPAGHAVGLFRAAV 1184
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               114 CGSSDLYLVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAV 173
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      54 ATQTFLATCINGVCWTVYHGAGTRIASPKGPVTQMYTNVDKDLVGWQAPQGSRSLTPCT 113
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              23
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              3 KKGSVVIVGRIN-----LSGDTAYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVST
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Genome polyprotein (Fragment).
Hepatitis C virus.
Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
Hepacivirus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SEQUENCE FROM N.A.
Choo O.-L., Richman K., Han J.;
Choo O.-L., Richman K., Han J.;
The nuclectide sequence of the Hepatitis C viral genome.";
Submitted (MAY-1990) to the EMBL/GenBank/DDBJ databases.
EMBL; M32084; AAA45677.1; -.
HSSP; P27958; IA1V.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      .4e-80;
les 13; indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   01-NOV-1996 (TrEMBLrel. 01, Created)
01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
01-WAR-2003 (TrEMBLrel. 23, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             PRT; 2436 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                86.1%; Score 888.5;
84.3%; Pred. No. 5.4e
tive 10; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     InterPro; IPR002518; HCV_NS1.

InterPro; IPR004109; HCV_NS3.

DR InterPro; IPR00149; HCV_NS4a.

DR InterPro; IPR00149; HCV_NS4b.

InterPro; IPR001668; HCV_NS4b.

R InterPro; IPR001669; HCV_NS4b.

R InterPro; IPR001669; HCV_NS4b.

InterPro; IPR001669; HCV_NS4b.

InterPro; IPR001699; RNA_PO1_PSv1r.

Pfam; PF01560; HCV_NS1; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      174 CTRGVAKAVDFIPVESLETIMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Best Local Similarity 84.3
Matches 172; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        PRELIMINARY;
```

```
[1]
SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 NCBI_TaxID-11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  STRAIN-colonel;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Q9ELS8;
Q9ELS8;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               RESULT 5
09ELS8
09ELS8
01. Mac O9ELS8
01. Mac O9ELS8
01. Mac O1. Mac O
            SOW WERE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   614
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              54 ATQTFLATCINGVCWTVYHGAGTRIIASPKGPVTQMYTNVDKDLVGWQAPQGSRSLTPCT 113
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      674
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              114 CGSSDLYLVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAV 173
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 R Pfam; PF012007; HCV_NS3; 1.

R Pfam; PF010001; HCV_NS4a; 1.

R Pfam; PF01001; HCV_NS4b; 1.

R Pfam; PF01051; HCV_NS5b; 1.

R Pfam; PF00271; HCV_NS5b; 1.

R Pfam; PF00271; HCV_NS5b; 1.

R Pfam; PF00271; HCV_NS1; 1.

R Probom; PD180662; HCV_NS1; 1.

R PROSTIE; PS50507; RDRP_POSITIVE; 1.

R PROSTIE; PS50507; RDRP_POSITIVE; 1.

R PROSTIE; PS50507; RDRP_POSITIVE; 1.

R HYP-binding; Coat protein; Envelope protein; Glycoprotein; Hclicasc; W Hydrolase; Nonstructural protein; Polyprotein; M. Hydrolase; Nonstructural protein; Polyprotein; M. Hydrolase; Nonstructural protein; Polyprotein; M. MA-directed RNA polymerase; Transferase; Transmembrane.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          3 KKGSVVIVGRIN-----LSGDTAYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVST
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           æ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  MEDLINE-21262212; PubMcd-11369872;
Lanford R.E., Lee H., Chavez D., Guerra B., Brasky K.M.;
"Infectious cDNA clone of the hepatitis C virus genotype 1 prototype
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               9; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Genome polyprotein.
Hepatitis C virus.
Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
Hepatulius.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  J. Geo. Virol. 82:1291-1297(2001).

-1- SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLECCAPSID COVERED BY LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
PROTEIN M AND GLYCOPROTEIN E. THE NUCLECCAPSID IS A COMPLEX OF PROTEIN C. AND MRNA (BY SIMILARITY).

EMBL. AF21632: AAF81759.1:
HSSP: P27958: 1A1V.
InterPro: IPR003445: Cytc_heme_bind.
InterPro: IPR001410; DEAD.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Length 2436;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Indels
                                                                                                                                                                                                                                                                                                                                                                                                                            2436 2436 24734 MW; D7B9872900BE3125 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Last sequence update)
Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Score 884.5; DB 12;
Pred. No. 1e-79;
9; Mismatches 14;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                PRT; 3011 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CTRGVAKAVDFIPVENLETIMRSP 758
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  174 CTRGVAKAVDFIPVESLETTMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  01-0CT-2000 (TrEMBLrel. 15, Created)
01-0CT-2000 (TrEMBLrel. 15, Last sequ
01-MAR-2003 (TrEMBLrel. 23, Last anno
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            HCV_capsid.
HCV_core.
HCV_env.
HCV_NSI.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  HCV_NS3.
HCV_NS4a.
HCV_NS4b.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             vuery Match 85.7%;
Best Local Similarity 84.3%;
Matches 172; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   IPR002518; H
IPR004109; H
IPR000745; E
IPR001490; E
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              InterPro; IPR002521;
InterPro; IPR002519;
InterPro; IPR002531;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  IPR002522;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          NCBI_TaxID-11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            sequence.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   InterPro;
InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              InterPro;
InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                               NON_TER
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                      NON_TER
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Q9IFE5
Q9IFE5;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     RESULT 4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              THE PART OF THE PA
            õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
```

```
1005 RRGREILLGPADGMVSKGWRLLAPITAYAQQTRGLLGCIITSLTGRDKNQVEGEVQIVST 1064
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    114 CGSSDLYLVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAV 173
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         3 KKGSVVIVGRIN-----LSGDIAYAQQIRGEEGCQETSQIGRDKNQVEGEVQIVSI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      54 ATQIFLATCINGVCWTVYHGAGTRIIASPKGPVTQMYINVDKDLVGWQAPQGSRSLTPCT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         9; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      PROSITE; PSO0190; CITCCHROME_C; 1.
PROSITE; PSS0190; RDRP_POSITIVE; 1.
PROSITE; PSS0507; RDRP_VIRAL; 1.
ATP-binding; Coat protein; Envelope protein; Glycoprotein; Helicase; Hydrolase; Monstructural protein; Polyprotein; RNA-directed RNA polymerase; Transferase; Transmembrane.
SEQUENCE 3011 AA; 327124 MW; 2489CE74AC864E58 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
Hepacivirus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               DB 12; Length 3011;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Ouery Match 85.7%; Score 884.5; DB 12; Length Best Local Similarity 84.3%; Pred. No. 1.4e-79; Matches 172; Conservative 9; Mismatches 14; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               01-MAR-2001 (TYEMBLrel. 16, Created)
1-MAR-2001 (TYEMBLrel. 16, Last sequence update)
01-MAR-2003 (TYEMBLrel. 23, Last annotation update)
Genome polyprotein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             PRT; 3011 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              InterPro; IPR002868; HCV_RGRB.
InterPro; IPR002166; HELICASEA.
InterPro; IPR001650; HELICASE_C.
InterPro; IPR007095; RNA_POA_DS_PS.
InterPro; IPR007095; RNA_POA_DS_PS.
InterPro; IPR007094; RNA_POA_DS_PS.
InterPro; IPR007094; RNA_POA_DS_PS.
InterPro; IPR007094; RNA_POA_DS_PS.
InterPro; IPR007094; RNA_POA_DS_PS.
IPR0000000001; HCV_COTE; I.
Pfam; PF01509; HCV_NSA; I.
Pfam; PF01001; HCV_NSA; I.
Pfam; PF00097; Viral_RRP; I.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                174 CTRGVAKAVDFIPVESLETTMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             PRELIMINARY;
```

7;

53

```
"The 5'-terminal sequence of the hepatitis C virus genome.";
Jpn. J. Exp. Med. 60:167-177(1990).
                                                                                                      MEDLINE-92044440; PubMed-1658196;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         HCV_RdRP.
Helicase_C.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               HCV_capsid.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     HCV_NS3.
HCV_NS4a.
HCV_NS4b.
HCV_NS5a.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 HCV_NS1.
HCV_NS2.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          HCV_env
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                EMBL; D10749; BAA01582.1; -. HSSP; P27958; 1HEI.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         InterPro; IPR001410; DEAD.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           InterPro; IPR002522; H
InterPro; IPR002512; H
InterPro; IPR002519; H
InterPro; IPR002518; H
InterPro; IPR004109; H
InterPro; IPR004409; H
InterPro; IPR001490; H
InterPro; IPR001490; H
InterPro; IPR001490; H
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         InterPro; IPR002166;
                                                                                                                                                                                                                                                                                SEQUENCE FROM N.A.
                                                             SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                   SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                            STRAIN-HC-J1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Inchauspe G.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    STRAIN-HC-J
                                                                                      STRAIN-HC-J
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Okamoto
SOW WE ARE THE TRANSPORT OF THE TRANSPOR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1065 ATQTFLATCINGVCMTVYHGAGIRTIASPKGPVIQMYINVDQDLVGWPAPQGSRSLTPCT 1124
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1125 CGSSDLYLVTRHADVIPVRRQGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGLFRAAV 1184
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ä
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    54 ATQTFLATCINGVCWTVYHGAGTRTIASPKGPVTQMYTNVDKDLVGWQAPQGSRSLTPCT 113
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          114 CGSSDLYLVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAV 173
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    3 KKGSVVIVGRIN------LSGDIAYAQQIRGEEGCQEISQIGRDKNQVEGEVQIVST
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       PROSITE; PS00190. CTYCCHROME_C; 1.
PROSITE; PS0507; RDRP_POSITIVE; 1.
ATP-DINGHING; COAL PRORE_UINAL; 1.
ATP-DINGHING; COAL PROTEUR: Brotelope protein; Glycoprotein; Helicase; Hydrolase; NonStructural protein; Polyprotein;
HYDrolase; NonStructural protein; Polyprotein;
HRNA-directed RNA polymerase; Transferase; Transmembrane.
SEQUENCE 3011 AA; 327107 MW; AGBECFSA3B3EE13F CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Hepatitis C virus.
Viruses: ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
Hepacivirus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   DB 12; Length 3011;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               DINALNE-91013116; PubMed-2170712;
Okamoto H., Okada S., Sugiyama Y., Yotsumoto S., Tanaka T.,
Yoshizawa H.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              85.7%; Score 884.5; DB 12; Length
83.8%; Pred. No. 1.4e-79;
ive 11; Mismatches 13; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Last sequence update)
Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             PRT; 3011 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1185 CTRGVAKAVDFIPVENLETTMRSP 1208
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            174 CTRGVAKAVDFIPVESLETTMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 01-NOV-1996 (TrEMBLrel. 01, Created)
01-NOV-1996 (TrEMBLrel. 01, Last seq
01-MAR-2003 (TrEMBLrel. 23, Last anno
                                                                                                                                                                                                         InterPro; IPR001650; Helicase_C.
InterPro; IPR007095; RNA_pol_DS_PS.
InterPro; IPR007094; RNA_pol_PSvir.
                                                                                                                                                                                                                                                                                                                                                               Pram; Pro1538; HCV_NS2; 1. Pram; Pr01538; HCV_NS2; 1. Pram; Pr01006; HCV_NS4; 1. Pram; Pr01001; HCV_NS4; 1. Pram; Pr01506; HCV_NS5a; 1. Pram; Pr00271; hellcase_C; 1. Pram; Pr00998; viral_RdRP; 1. Pr0Dom; PD186062; HCV_NS1; 1.
                                           HCV_env.
HCV_NS1.
HCV_NS3.
HCV_NS3.
HCV_NS4.
HCV_NS4B.
HCV_NS5B.
                                                                                                                                                                                                                                                                                  PF01543; HCV_capsid;
                                                                                                                                                                                                                                                                                                        HCV_core;
HCV_env; 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Local Similarity 83.8
hes 171; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             PRELIMINARY:
                                                                                                                                                                                                                                                                                                                                                 PF01560; HCV_NS1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SM00487; DEXDC;
                                                             InterPro; IPR002531; H
InterPro; IPR002518; H
InterPro; IPR004109; H
InterPro; IPR000745; H
InterPro; IPR001490; H
                  IPR002521;
IPR002519;
                                                                                                                                                                          InterPro; IPR002868;
                                                                                                                                                                                              IPR002166;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SEQUENCE FROM N.A. STRAIN-HC-J1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         NCBI_TaxID-11103;
                                                                                                                                                                                                  InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SMART; SM
PROSITE;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          003463
                                                                                                                                                                                                                                                                                  Pfam;
                                                                                                                                                                                                                                                                                                                             Pfam;
Pfam;
Pfam;
Pfam;
Pfam;
Pfam;
Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Best Loca
Matches
                                                                                                                                                                                                                                                                                                      Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  RESULT 6
```

a

a

ð

qa ð

```
Okamoto H., Okada S., Sugiyama Y., Kurai R., Iizuka H., Machida A., Miyakawa Y., Mayumi M.;
"Nucleotide sequences of the genomic RNA of hepatitis C virus isolated from a human carrier: comparison with reported isolates for conserved and divergent regions.";
J., Gen. Virol. 72:2697-2704(1991).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ATP-binding; Coat protein; Envelope protein; Glycoprotein; Helicase;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Okamoto H., Kanai N., Mishiro S.; "Full-length nucleotide sequence of a Japanese hepatitis C virus "Full-length nucleotide sequence of isolate (HC-JI) with high homology to USA isolates."; Nucleic Acids Res. 20:6410-6410
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           MEDLINE-94174722; PubMed-7510436;
Mink M., Benichou S., Madaule P., Tiollais P., Prince A.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Hydrolase; Nonstructural protein; Polyprotein;
RNA-directed RNA polymerase; Transferase; Transmembrane.
SEQUENCE 3011 AA; 327112 MW; 97E9052C0250463B CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Submitted (DEC-1992) to the EMBL/GenBank/DDBJ databases
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            InterPro: IPR001650; Helicaser.
InterPro: IPR001650; RALpol_DS_PS.
InterPro: IPR001094; RNA_pol_DS_PS.
InterPro: IPR007094; RNA_pol_DS_PS.
InterPro: IPR007094; RNA_pol_DS_PS.
InterPro: IPR007094; RNA_pol_DS_PS.
InterPro: IPR007094; RNA_pol_DS_PS_IPR007091; RNA_pol_DS_PS_IPR00709099; RNA_PS_IPR0070999; VIRAL_RGRP; I.
Pfam: PP001001; HCV_NS4b; I.
Pfam: PP001001; HCV_NS4b; I.
Pfam: PP001001; HCV_NS4b; I.
Pfam: PP00101; HCV_NS5a; I.
Pfam: PP00101; HCV_NS5a; I.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SMART; SM00487; DEXDC; 1.
PROSITE; PS50507; RDRP_POSITIVE; 1.
PROSITE; PS50521; RDRP_VIRAL; 1.
                                                                                                                                                                                                                                                                                STRAIN-HC-J1;
MEDLINE-93117120; PubMed-1335573;
```

```
HSSP; P27958; 1HEI.
InterPro; IPR001410;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              IPR002531;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  IPR002868;
IPR002166;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    IPR002518;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         IPR004109;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             IPR001490;
                                                                                                                                                                                                                           Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               NCBI_TaxID-11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  InterPro;
InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          InterPro;
InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           chimeras.
                                                                                                                                                                                                        Query Match
                                                                                                                                                                                                                           Local
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Q9PWX5
Q9PWX5;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          RESULT 8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Q9PWX5
DR KKE DR K KE
                                                                                                                                                                                                                                                                                                                         qq
                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                          qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ,,
                                                                                                                                                                                                                                                                                     1125 CGSSDLYLVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHVVGIFRAAV 1184
                                                                                                                                        1005 RKGREILLGPADGMYSKGWRLLAPITAYAQQTRGLLGCIITSLTGRDKNQVEGEVQIVST 1064
                                                                                                                                                                                                                         1065 AAQTFLATCINGVCWTVYHGAGTRTIASPKGPVIQMYTNVDQDLVGWPAPQGARSLTPCT 1124
                                                                                                                                                                                114 CGSSDLYLVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAV 173
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    C SUBMIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A LIPPOPROTEIN ENVELOPE CONSISTS OF TWO PROTEINS:

LIPOPAGENTH: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A LIPPOPROTEIN ENVELOPE CONSISTS OF TWO PROTEINS:

PROTEIN AND GLYCOPPORTEN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MRNA (BY SIMILARITY).

EMBL: AFOLITS!; AAB67036.1; -

REMEL: AFOLITS!; AAB67036.1; -

RICEPPO: IPRO01410; DEAD.

RICEPPO: IPRO02521; HCV_core.

RICEPPO: IPRO02521; HCV_core.

RICEPPO: IPRO02519; HCV_NS2.

RICEPPO: IPRO02519; HCV_NS2.

RICEPPO: IPRO04109; HCV_NS3.

RICEPPO: IPRO04109; HCV_NS3.

RICEPPO: IPRO04096; HCV_NS4.

RICEPPO: IPRO04096; HCV_NS4.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Yanagi M., Purcell R.H., Emerson S.U., Bukh J.;
"Transcripts from a single full-length cDNA clone of hepatitis C virus are infectious when directly transfected into the liver of a
                                                                                                  3 KKGSVVIVGRIN-----LSGDTAYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVST
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    BY A
                                                           Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Genome polyprotein.
Beparitis C virus strain H77.
Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
Hepacivirus
                                                         6
                 DB 12; Length 3011;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               01-JAN-1998 (TrEMBLrel. 05, Created)
01-JAN-1998 (TrEMBLrcl. 05, Last sequence update)
01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
              Score 883.5; DB 12;
Pred. No. 1.7e-79;
8; Mismatches 15;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           PRT; 3011 AA
                                                                                                                                                                                                                                                                                                                                                                      1185 CTRGVAKAVDFIPVESLETTMRSP 1208
                                                                                                                                                                                                                                                                                                                                              174 CTRGVAKAVDFIPVESLETTMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   IPR001650; Helicase_C.
IPR007095; RNA_pol_DS_PS.
IPR007094; RNA_pol_PSvir.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     STRAIN-H77;
MEDLINE-97385173; PubMed-9238047;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              HCV_capsid; 1.
HCV_core; 1.
HCV_env; 1.
HCV_NS1; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  HCV_RdRP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            HCV_NS5a; 1.
helicase_C; 1
              85.68;
84.38;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               HCV_NS2; 1.
HCV_NS3; 1.
HCV_NS4a; 1.
HCV_NS4b; 1.
                                      Best Local Similarity 84.3
Matches 172; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  InterPro; IPR002166;
InterPro; IPR001650;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        InterPro; IPR007095;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              NCBI_TaxID-63746;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  PF01506;
PF00271;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                PF01543;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      PF01542;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         PF01539;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 PF01538;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            PF01006;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               PF01001;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            chimpanzee.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       interPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             InterPro;
                   Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           036608
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                RESULT 7
036608
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                DER RESERVED DER PRESERVEN DER
                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                      a
                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                   ö
```

```
114 CGSSDLYLVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAV 173
                                                                                                                                                                                                                                                                                                                                                                                          53
                                                                                                                                                                                                                                                                                                                                                                                          3 KKGSVVIVGRIN-----LSGDTAYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVST
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  54 ATQTFLATCINGVCWTVYHGAGTRTIASPKGPVTQMYTNVDKDLVGWQAPQGSRSLTPCT
Pfam; PF00998; Viral_RdRP; 1.
ProDom: PD186062; HCV_NS1; 1.
SMART: SMOGH87; DEXDC; 1.
PROSITE; PS50501; RDRP_POSITIVE; 1.
PROSITE; PS50521; RDRP_VIRAL; 1.
ATP-binding; Coat protein; Envelope protein; Glycoprotein; Helicase; Hydrolase; Monstructural protein; Polyprotein; BNA-directed RNA polymerase; Transferase; Transmembrane.
SEQUENCE 3011 AA; 327112 MM; 0B75E6B81CB5C198 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  K
                                                                                                                                                                                                                                                                                                                                       9; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SEQUENCE FROM N.A.

SEQUENCE FROM N.A.

PEDELINE-94203956, PubMed-10489358;

Yanagi M., Purcell R.H., Emerson S.U., Bukh J.;

"Hepatitis C virus: an infectious molecular clone of a second major manage (2a) and lack of viability of intertypic la and 2a chimers ".
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Bukh J.;
Submitted (AUG-1999) to the EMBL/GenBank/DDBJ databases.
--- SUBGNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY
LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF
PROTEIN C AND MRNA (BY SIMILARITY).
EMBL; AF177040; AAF01180.1; --
EMBL; AF177040; AAF01180.1; --
                                                                                                                                                                                                                                                                                 DB 12; Length 3011;
                                                                                                                                                                                                                                                                                                                                       14; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             01-NAY-2000 (TrEMBLrel. 13, Created)
01-NAY-2000 (TrEMBLrel. 13, Last sequence update)
01-NAR-2003 (TrEMBLrel. 23, Last annotation update)
Genome polyprotein.
                                                                                                                                                                                                                                                                                 85.4%; Score 881.5; DB 1:
83.8%; Pred. No. 2.7e-79;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                PRT; 3015 AA.
                                                                                                                                                                                                                                                                                                                                    10; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1185 CTRGVAKAVDFIPVENLGTTMRSP 1208
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     174 CTRGVAKAVDFIPVESLETIMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                HCV_capsid.
HCV_core.
HCV_env.
HCV_NSI.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              HCV_NS4a.
HCV_NS4b.
HCV_NS5a.
HCV_RdRP.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           HCV_NS2.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Virology 262:250-263(1999)
                                                                                                                                                                                                                                                                                                                                 Matches 171; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                PRELIMINARY;
```

```
PROTEIN C AND MRNA (BY SIMILARITY).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Hepacivirus.
NCBI_TaxID-11103;
                                                                                                                      P27958;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             091RR8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     091RR8
                                           EMBL;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      RESULT 10
Q91RR8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SOW WENT WAS A STREET OF THE S
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ô
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ä
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   54 ATQTFLATCINGVCWTVYHGAGTRT1ASPKGPVTQMYINVDKDLVGWQAPQGSRSLTPCT 113
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 114 CGSSDLYLVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAV 173
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 53
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 3 KKGSVVIVGRIN-----LSGDTAYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVST
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Bukh J.;
Submitted (AUG-1999) to the EMBL/GenBank/DDBJ databases.
-!- SUBGNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A
-!- SUBGNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID FOVERED BY A
-!- EUPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  9; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              PROSITE; PSO0392: DDC_GAD_HDC_XDC; 1.
PROSITE; PSO50507: RDRP_POSITIVE; 1.
ATP-SIAGNOST PRORP_VIRAL; 1.
ATP-SIAGNOST CONTROL OF PROSITE; PSO5021; PRORP_VIRAL; 1.
ATP-SIAGNOST CONSTRUCTURAL PROTEIN; FOUND PROPERTH; FOUND REPROSITE; PROSITE OF 
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 MEDIINE-99420196; PubMed-10489358;
Yanaqi M., Purcell R.H., Emerson S.U., Bukh J.;
"Hepatitis C virus: an infectious molecular clone of a second major
genotype (za) and lack of viability of intertypic la and 2a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Genome polyprotein.

Hepatitis C virus.
Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
Hepacivirus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               DB 12; Length 3015;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               85.4%; Score 881.5; DB 12; Length
83.8%; Pred. No. 2.7e-79;
Live 10; Mismatches 14; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           01-WAY-2000 (TrEMBLrcl. 13, Created)
01-WAY-2000 (TrEMBLrcl. 13, Last sequence update)
01-WAR-2003 (TrEMBLrcl. 23, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                PRT; 3015 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1189 CTRGVAKAVDFIPVENLGTTMRSP 1212
                          InterPro; IPR00120; Piridoxal_dec.
InterPro; IPR00120; RNA_pol_DS_PS.
InterPro; IPR007095; RNA_pol_DS_PS.
InterPro; IPR007095; RNA_pol_DS_PS.
InterPro; IPR007094; RNA_pol_DS_PS.
InterPro; IPR007094; RNA_pol_DS_PS.
IPfam; PP01542; HCV_core; 1.
Pfam; PP01563; HCV_core; 1.
Pfam; PP01563; HCV_NS1; 1.
Pfam; PP01006; HCV_NS3; 1.
Pfam; PP01006; HCV_NS4; 1.
Pfam; PP01001; HCV_NS4s; 1.
Pfam; PP01001; HCV_NS4s; 1.
Pfam; PP01001; HCV_NS4s; 1.
Pfam; PP01001; HCV_NS4s; 1.
Pfam; PP01001; HCV_NS5s; 1.
Pfam; PP00271; helicase 2.
Pfam; PP00271; helicase 2.
Ppr00070; PD186062; HCV_NS1; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   174 CTRGVAKAVDFIPVESLETTMRSP 197
IPR001650; Helicase_C
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Virology 262:250-263(1999).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Query Match 85.4 Best Local Similarity 83.8 Matches 171; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SMART; SM00487; DEXDC; 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   NCBI_TaxID-11103;
    InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             chimeras.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            09PW09;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                09PW09
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 RESULT 9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         6DMd6C
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         HERE REPRESENTED TO THE PROPERTY OF THE PROPER
g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    윱
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
```

```
54 ATQTFLATCINGVCWTVYHGAGTRTIASPKGPVTQMYTNVDKDLVGWQAPQGSRSLTPCT 113
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    114 CGSSDLYLVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAV 173
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    3 KKGSVVIVGRIN------LSGDTAYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVST
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Helicase;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     85.4%; Score 881.5; DB 12; Length 3015;
83.8%; Pred. No. 2.7e-79;
tive 10; Mismatches 14; Indels 9;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   PROSITE; PS00392; DDC_GAD_HDC_YDC; 1.
PROSITE; PS50507; RDRP_POSITIVE; 1.
PROSITE; PS50507; RDRP_VIRAL; 1.
ATP-Dinding; Coat protein; Envelope protein; Glycoprotein; Hydrolase; Nonstructural protein; Polyprotein;
RNA-directed RNA polymerase; Transferase; Transmembrane.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          3015 AA; 328084 MW; E309F6318067D6CD CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Created)
Last sequence update)
Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       181 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     174 CTRGVAKAVDFIPVESLETTMRSP 197
                                                                                                                                                                                                                                                                                                                                                                      InterPro; IPRO01029; PFITGORAL JOEC InterPro; IPRO01029; RNA_pol_DS_PS. InterPro; IPR001094; RNA_pol_DS_PS. InterPro; IPR001094; RNA_pol_DS_PS. InterPro; IPR001094; RNA_pol_DS_PS. InterPro; IPR001054; RNA_pol_DS_PS. InterPro; IPR001050; RCV_core; I. Pfam; PP01550; RCV_RS; I. Pfam; PP01550; RCV_RS; I. Pfam; PP01006; RCV_RS; I. Pfam; PP01006; HCV_RS3; I. Pfam; PP01001; HCV_RS4; I. Pfam; PP0100998; Viral_RGRP; I. PR000998; Viral_RGRP; I. PR000001; PD16662; RCV_RS1; I.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          PRT;
                                                                                                 HCV_capsid.
HCV_core.
                                                                                                                                                    HCV_env.
HCV_NS1.
HCV_NS2.
HCV_NS3.
HCV_NS4a.
HCV_NS5a.
HCV_RS5a.
HCV_RGRP.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     01-DEC-2001 (TEMBLEEL 19, 01-DEC-2001 (TEMBLEEL 19, 01-MR-2003) (TEMBLEEL 23, NG3) protease (Fragment). Hepatitis C virus.
AF177039; AAF01181.1;
AF177037; AAF01179.1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Best Local Similarity 83.8 Matches 171; Conservative
                                                                           InterPro; IPR001410; DEAD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          PRELIMINARY;
                                                                                                                                               Interpro; IPR002519; H
Interpro; IPR002519; H
Interpro; IPR002518; H
Interpro; IPR004109; H
Interpro; IPR001459; H
Interpro; IPR001456; H
Interpro; IPR002868; H
Interpro; IPR002166; H
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                DEXDC;
                                                                                                 InterPro; IPR002522;
InterPro; IPR002521;
```

a

ò

ŏ

a à

```
79 IASPRGPVTQMYTNVDKDLVGWQAPQGSRSLTPCTCGSSDLYLVTRHADVIPVRRRGDSR 138
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             4 TAYAQQTRGLLGCIXTSLTGRDKNQVEGEVQIVSTAAQTFLATCINGVCWTVYHGAGTRT 63
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           19 TAYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVSTATQTFLATCINGVCWTVYHGAGTRT 78
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            124 GSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVCTRGVAKAVDFIPVENLETTMRS 181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Holland-Staley C.A., Kovari L.C., Golenberg E., Mayers D.L.; "Genetic Diversity and response to IFN of the NS3 Protease Gene from Clinical Strains of the Hepatitis C Virus.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Holland-Staley C.A., Kovari L.C., Golenberg E., Mayers D.L.;
"Genetic Diversity and response to IFN of the NS3 Protease Gene from
"Clinical Strains of the Hepatitis C Virus.";
Submitted (APR-2001) to the EMBL/GenBank/DDBJ databases.
EMBL, AF865238, AR8545633.1;
Interpro: IPR004109; HCV_NS3.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
Hepacivirus.
                                                                                                                                                                                                                                                                                                                                                                                                                                            Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ;
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Length 181;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   8; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Submitted (APR-2001) to the EMBL/GenBank/DDBJ databases EMBL, AR186241; AR18646.1; -. InterPro; IRR04109; HCV_NS3. Pfam: PF02907; HCV_NS3; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       19123 MW; 1CAE817345ED809D CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 181 181
181 Aa; 19084 MW; 3B5E8161F2100A72 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Last sequence update)
Last annotation update)
                                                                                                                                                                                                                                                                                            Last sequence update)
Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       85.0%; Score 877; DB 12;
94.4%; Pred. No. 2.1e-80;
tive 2; Mismatches 8;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              181 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       01-0EC-2001 (TrEMBLrel. 19, Created)
01-0EC-2001 (TrEMBLrel. 19, Last sequ
01-MAR-2003 (TrEMBLrel. 23, Last anno
NS3 protease (Fragment).
Hepatitis C virus.
                                                                                                                                                                                                                                                          Created)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           PRT;
                                                                                                                                                                                                                                                                                            01-DEC-2001 (TrEMBLrel. 19, 01-MAR-2003 (TrEMBLrel. 23,
                                                                                                                                                                                                                                                      01-DEC-2001 (TrEMBLrel. 19,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Best Local Similarity 94.4
Matches 168; Conservative
                                                                                                                                                                                                                                                                                                                                                                    NS3 protease (Fragment).
Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           PRELIMINARY;
                                                                                                                                                                                  PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 181 1
181 AA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      NCBI_TaxID-11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         NCBI_TaxID-11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                STRAIN-Pt. 4V
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               STRAIN-Pt. 3U;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           NON_TER
NON_TER
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  NON_TER
NON_TER
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Protease.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Protease.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    139
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   RESULT 13
                                                                                                                                            091RR5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ACCOUNT REPORT R
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
g
                                                                                                                                                                                      SO SET THE SET OF SO SET THE SO SET THE SET OF SET THE SET THE SO SET THE SET THE SO SET THE S
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    79 IASPKGPUTOMYTNVDKDLVGWQAPQGSRSL/PCTCGSSDLYLVTRHADVIPVRRRGDSR 138 [[[[[[[[]]]]]]]]]]]]
64 IASPKGPVIQMYTNVDKDLVGWPAPQGSRSL/PCTCGSSDLYLVTRHADVIPVRRRGDSR 123
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             79 IASPKGPVTQMYTNVDKDLVGWQAPQGSRSLTPCTCGSSDLYLVTRHADVIPVRRRGDSR 138
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       64 IASPKGPVIQMYINVDKDLVGWPAPQGSRSLTPCTCGSSDLYIVTRHADVIPVRRRGDSR 123
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                19 TAYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVSTATQTFLATCINGVCWTVYHGAGTRT 78
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      4 TAYAQQTRGLLGCIITSLTGRDKNQVEGEVQIVSTAAQTFLATCINGVCWTVYHGAGTRT 63
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 4 TAYAQOTRGLLGCIITSLTGRDKNOVEGEVQIVSTAAQTFLATCINGVCWTVYHGAGTRT 63
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               19 TAYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVSTATQTFLATCINGVCWTVYHGAGTRT 78
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             139 GSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVCTRGVAKAVDFIPVESLETIMRS 196
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          139 GSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVCTRGVAKAVDFIPVESLETTMRS 196
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Holland-Staley C.A., Kovari L.C., Golenberg E., Mayers D.L.;
"Genetic Diversity and response to Ifro of the NS3 Protease Gene from
"Clinical Strains of the Hepatitis C Virus.";
Submitted (APR-2001) to the EMBL/GenBank/DDBJ databases.
EMBL, AF8654918, AAR5443.1:
InterPro; IPR004109; HCV.NS3.
PF02907; HCV_NS3; I.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Gaps
                                                                    STRAIN-Pt.1Y;
Holland-Staley C.A., Kovari L.C., Golenberg E., Mayers D.L.;
"Genetic Diversity and response to IFN of the NS3 Protease Gene from Clinical Strains of the Hepatitis C Virus.';
Submitted (APR-2001) to the EMBL/GenBank/DDBJ databases.
EMBL; AF369235; AAK54560.1;
InterPro; IPR004109; HCV_NS3.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Length 181;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Length 181;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Score 879; DB 12; Length 19
Pred. No. 1.3e-80;
1; Mismatches 8; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            181 AA; 19130 MW; 85D91869299B7C35 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                     181 181
181 AA; 19130 MW; 85D91869299B7C35 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Q91RT5;
U-DEC-2001 (TrEMBLrel. 19, Created)
01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Score 879; DB 12;
Pred. No. 1.3c-80;
1; Mismatches 8;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       181 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               85.2%;
94.9%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    tch 85.2%; al Similarity 94.9%; 169; Conservative
                                                                                                                                                                                                                                                                                                                           Pfam; PF02907; HCV_NS3; 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       169; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      NS3 protease (Fragment).
Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Query Match
Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        [1]
SEQUENCE FROM N.A.
                                     SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Hepacivirus.
NCBI_TaxID-11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                STRAIN-Pt. 4;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        NON_TER
NON_TER
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Protease.
                                                                                                                                                                                                                                                                                                                                                                                                                                            NON_TER
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Query Match
                                                                                                                                                                                                                                                                                                                                                                                                      NON_TER
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Q91RT5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Best Loca
Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Best Loca
Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   RESULT 11
091RT5
```

g ò a

ö

```
SEQUENCE FROM N.A.
NCBI_TaxID=11103;
                                                                                                                                                                                                                                                                                                                                                         NON_TER
SEQUENCE
                                                                                                                                                                                                                                                                                                     Protease.
NON_TER
                                                                                                                                                                                                                                                                                                                                                                                                                                      Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Best Loca
Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Search cor
Job time
     OX REPRESENT STATE OX SOLUTION OF THE STATE OF THE SOLUTION OF
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ö
                                                                                   ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             79 IASPKGPVTOMYINVDKDLVGWQAPQGSRSLTPCTCGSSDLYLVTRHADVIPVRRRGDSR 138
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             64 IASPKGPVIQMYTNVDKDLVGWPAPQGARSLTPCTCGSSDLYLVTRHADVIPVRRRGDSR 123
                                                                                                                                                                                                                                                                           64 IASPKGPVIQMYINVDQDLVGWPAPQGARSLTPCTCGSSDLYLVTRHADVIPVRRRGDSR 123
                                                                                                                                                                                                                                              79 IASPRGPVTQMYTNVDKDLVGWQAPQGSRSLTPCTCGSSDLYLVTRHADVIPVRRRGDSR 138
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         19 TAYAQOTRGEEGCQETSQTGRDKNQVEGEVQIVSTATQTFLATCINGVCWTVYHGAGTRT 78
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     4 TAYAQOTRGLLGCIITSLTGRDKNQVEGEVQIVSTATQTFLATCINGVCWTVYHGAGTRI 63
                                                                                                                                                                    4 TAYAQQTRGLLGCIXTSLTGRDKNQVEGEVQIVSTAXQTFLATCINGVCWTVYHGAGTRT 63
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           GSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVCTRGVAKAVDFIPVESLETTMRS 196
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       124 GGLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVCTRGVAKAVDFIPVENLETTMRS 181
                                                                                                                                                                                                                                                                                                                                                      GSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVCTRGVAKAVDFIPVESLETTMRS 196
                                                                                                                                                                                                                                                                                                                                                                                   19 TAYAQOTRGEEGCQETSQTGRDKNQVEGEVQIVSTATQTFLATCINGVCWTVYHGAGTRT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Holland-Stalley C.A., Kovari L.C., Golenberg E., Mayers D.L.;
"Genetic Diversity and response to IFN of the NS3 Protease Gene from Clinical Strains of the Hepatitis C Virus.";
Submitted (APR-2001) to the EMBL/GenBank/DDBJ databases.
EMBL; AF369214: AAK54539.1;
Interpro: IPR004109; HCV_NS3.
Protease.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Gaps
                                                                                      Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Q91RR3 PRELIMINARY; PRT; 181 AA.

Q91RR3;
01-DEC-2001 (TFEMBLrel. 19, Created)
01-DEC-2001 (TFEMBLrel. 19, Last sequence update)
01-MAR-2003 (TFEMBLrel. 23, Last annotation update)
101-MAR-2003 (TFEMBLrel. 23, Last annotation update)
Hepatitis C virus.
Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae; Hepacivirus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  NS3 protease (Fragment).
Hepatitis C virus.
Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        0;
                                                                                   ;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              85.0%; Score 877; DB 12; Length 181; 94.4%; Pred. No. 2.1e-80; tive 2; Mismatches 8; Indels
                         Length 181;
                                                                                Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                181 181
181 AA; 19114 MW; BEIDOB542F014F86 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  01-DEC-2001 (TrEMBLrel. 19, Created)
01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
                                                                                   .;
8
                    Score 877; DB 12;
Pred. No. 2.1e-80;
2; Mismatches 8;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    181 AA
                       85.08;
94.48;
                       Query Match
Best Local Similarity 94.49
Matches 168; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Query Match 85.0°
Best Local Similarity 94.4°
Matches 168; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             NCBI_TaxID-11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Hepacivirus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         NON_TER
NON_TER
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                      139 (
                                                                                                                                                                                                                                                                                                                                                                                                          124
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              091RT9
091RT9;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            RESULT 15
091RR3
1D 091RR
AC 091RR
DT 01-DE
DT 01-DE
DE NS3 P
OS HEPAT
OC HEPAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             RESULT 14
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         q
                                                                                                                                          ö
                                                                                                                                                                                        g
                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                               a
                                                                                                                                                                                                                                                                                                                                                         ö
                                                                                                                                                                                                                                                                                                                                                                                                          q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ద
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           õ
```

```
79 IASPKGPVTQMYTNVDKDLVGWQAPQGSRSLIPCTCGSSDLYLVTRHADVIPVRRRGDSR 138
                                                                                                                                                                                                                                                                                                                                                                                                                     139 GSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVCTRGVAKAVDFIPVESLETTMRS 196
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  STRAIN=Pt.4B;
Holland-Staley C.A., Kovari L.C., Golenberg E., Mayers D.L.;
"Genetic Diversity and response to IFN of the NS3 Protease Gene from Clinical Strains of the Hepatitis C Virus.";
Submitted (APR-2001) to the EMBL/GenBank/DDBJ databases.
EMBL, RASGA540, RARA5455.1;
InterPro; IPR004109; HCV_NS3.
Pfam; PF02907; HCV_NS3; I.
                                                                                                                                                                                                                                                                                                                                                                              Gaps
                                                                                                                                                                                                                                                                                                                                                                         0;
                                                                                                                                                                                                                                                                                                                     84.9%; Score 876; DB 12; Length 181; larity 94.4%; Pred. No. 2.6e-80; Conservative 2; Mismatches 8; Indels
                                                                                                                                                                                                                                              181 181
181 AA; 19115 MW; 5D85F88AD7AC1A11 CRC64;
                                                                                                                                                                                                                                                                                                                                               Local Similarity
Les 168; Conserv
```

ö

completed: August 30, 2003, 19:18:19 e : 37.5921 secs

~

```
AR145264 Sequence
AR145263 Sequence
AR145263 Sequence
AR145263 Sequence
AR145265 Sequence
AR145267 Sequence
AR145269 Sequence
AR145269 Sequence
AR145263 Sequence
AR145263 Sequence
AR145263 Sequence
AR145263 Sequence
AR145263 Sequence
AR145273 Sequence
BD081910 Hepatitis
AR10808 Sequence
BD069982 Functiona
AR118686 Sequence
BD069985 Functiona
AR118686 Sequence
BD069985 Functiona
AR118686 Sequence
BD069385 Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  106440 Sequence 54
109329 Sequence 10
AR118696 Sequence
109311 Sequence 15
M22084 Hepatitis C
AR118703 Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          AR118692 Sequence
                                                                                                                                           Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.
                                                                                                                                                                                                                       Description
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ALIGNMENTS
                                                                                                                                                                                       SUMMARIES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         AR127810
BD081911
AR127809
BD081910
AR110828
BD069982
AR110831
BD069985
AR118686
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                HPCPOLYP
AR118703
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             109328
AR118692
                                                                                                                                                                                                                                                                                                                                                                                   ARI45253
ARI45265
ARI45274
ARI45272
ARI45250
ARI45250
ARI45250
ARI45251
ARI45251
ARI45271
ARI45271
ARI45271
ARI45271
                                                                                                                                                                                                                                                                                                                                        AR145269
                                                                                                                                                                                                                                                                                                                                                                        AR145252
                                                                                                                        em_htgo_other:*
             em_htg_inv:*
em_htg_other:*
                                            em_htg_pln:*
em_htg_rod:*
em_htg_mam:*
em_htg_wrt:*
                                                                                        em_sy:*
em_htgo_hum:*
                                   em_htg_mus:*
    em_htg_hum:*
                                                                                                                                                                                                          Query
Match Length DB
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        9646
12980
12980
5360
5360
                                                                                                                                                                                                                                         9922.5
9903.5
9003.5
904.5
904.5
904.5
903.5
900.5
900.5
900.5
900.5
900.5
900.5
900.5
                                                                                                                                                                                                                                                                                                                                                                                                                                     897.5
897.5
897.5
896.5
896
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    893.5
893.5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  \mathsf{HHILLHHIII}
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               RESULT
                                                                                                                                                                                                           Result
                                                                                                                                                                                                                      ģ
                                                                                                                                                                                                                                                                                                                                 August 30, 2003, 19:18:33 ; Search time 2560.57 Seconds (without alignments) 3147.423 Million cell updates/sec
                                                                                                                               1 MKKKGSVVIVGRINLSGDTA.....VAKAVDFIPVESLETTMRSP 197
                                                                                                                                                                                                                                         5777422
GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.
                                           using frame_plus_p2n model
                                                                                                                                                                                                                                      Total number of hits satisfying chosen parameters:
                                                                                                                                                                                                                   2888711 segs, 20454813386 residues
                                                                                                                                                                                                                                                                                          Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries
                                                                                                                                                              0.5
                                                                                                                                                           Xgapop 10.0 , Xgapext (Ygapop 10.0 , Ygapext (Fgapop 6.0 , Fgapext Delop 6.0 , Delext
                                                                                                                                                                                                                                                            Minimum DB seq length: 0
Maximum DB seq length: 2000000000
                                                                                                                                                                                                                                                                                                                                                                                                                                                         GenEmbl:*

1: gb_ba:*

2: gb_htg:*

3: gb_ni:*

4: gb_ow:*

9b_ph:*

9b_ph:*

9b_ph:*

9b_ph:*

9b_pr:*

9b_pr:*
                                                                                                         US-09-965-594-16
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            em_fun:*
em_hum:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               em_in:
em_mu:
em_om:
em_or:
em_ov:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 em_sts:*
em_un:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 gb_sy:*
gb_un:*
gb_vi:*
em_ba:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 em_ph:*
em_pl:*
                                           - nucleic search,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        em_ro:*
                                                                                                                                                    BLOSUM62
                                                                                                                    1032
                                                                                                         Title:
Perfect score:
                                                                                                                                                    Scoring table:
                                           OM protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Database :
                                                                                                                                Sequence:
                                                                                                                                                                                                                   Searched:
                                                                Run on:
```

```
Unclassified
AR145264
AR145264.1
                                                                                                                                                                  Percent Similarity:
Best Local Similarity:
Query Match:
DB:
                                                                                                                    411 a
                      Unknown
                                  Unknown
                                                                                                                                             Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                  122
                                                                                                                                                                                                                                                                                                                                                                                                                                                  424
                                                                                                                                                                                                                                                                                                                                 62
                                                                                                                                                                                                                                                                                                                                                 244
                                                                                                                                                                                                                                                                                                                                                                                  304
                                                                                                                                                                                                                                                                                                                                                                  82
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               DEFINITION
ACCESSION
VERSION
                                                                                                     source
                                                                                                                   BASE COUNT
ORIGIN
                                  ORGANISM
                                                                                                                                                    Pred. No.:
ACCESSION
VERSION
KEYWORDS
SOURCE
                                                          AUTHORS
TITLE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       RESULT 3
AR145268
                                                  REFERENCE
                                                                                   JOURNAL
                                                                                            FEATURES
                                                                                                                                                                                                                                                                                                                                                                                 a
                                                                                                                                                                                                                                                                                                                                                                                                  ð
                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                  a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      셤
                                                                                                                                                                                                                                           a
                                                                                                                                                                                                                                                            õ
                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                            셤
                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                             q
                                                                                                                                                                                                                                                                                                                                                                Š
                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         PAT 08-AUG-2001
                                                                                                                                                                                                                                                                                                                                                                                                                 20-APR-2002
                                                                                                                                                                                                                                                                                                                                                                                          532
                                                                                                                                                                                                                                                                                                                                 592
                                                                                                                                                                                                                                                                                                                                                                  652
                                                                                                                                                                                                                                                                                                                                                                                                                                                   ValThrArgHisAlaAspValIleProValArgArgArgGlyAspSerArgGlySerLeu 141
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     AlaGlyHisAlaValGlyIlePheArgAlaAlaValCysThrArgGlyValAlaLysAla 181
                                                                                                                                                                                                                                                                                                                                                                                  ProLysGlyProValThrGlnMetTyrThrAsnValAspLysAspLeuValGlyTrpGln 101
                                                                                                                                                                                                                                                                                                                                                  81
                                                                                                                                                                                                                                                                                41
                                                                                                                                                                                                                                                                                                                61
                                                                                                                                                                                                                                                21
                                                                                                                                                                                                                                                                              AlaGlnGlnThrArgGlyGluGluGlyCysGlnGluThrSerGlnThrGlyArgAspLys
                                                                                                                                                                                                                                                                                         GlySerValValIleValGlyArgIleAsnLeuSerGlyAsp-----ThrAlaTyr
                                                                                                                                                                                                                                                                                                                AsnGlnValGluGlyGluValGlnIleValSerThrAlaThrGlnThrPheLeuAlaThr
                                                                                                                                                                                                                                                                                                                                                  CyslleAsnGlyValCysTrpThrValTyrHisGlyAlaGlyThrArgThrIleAlaSer
                                                                                             pestivirus
         PAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ValAspPhelleProValGluSerLeuGluThrThrMetArgSer 196
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         linear
         linear
                                                                  Unclassified.

1 (bases 1 to 12734)

Hong,Z., Lai,V.C.H. and Lau,J.Y.N.

Hong,Z. Lai,V.C.H. and Lau,J.Y.N.

Papelitis C virus protease-dependent chimeric patent: US 6326137-A 1 04-DEC-2001;
                                                                                                                                                                    Length:
Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        AR145264 1998 bp I Sequence 105 from patent US 6211338.
                                                                                                                                      2803
      Sequence 1 from patent US 6326137.
AR179057 GI:20220612
                                                                                                                                                                                                                               US-09-965-594-16 (1-197) x AR179057 (1-12734)
                                                                                                                                                                                                              Gaps:
                                                                                                                          /organism-"unknown"
1 2604 c 3295 g
                                                                                                            Location/Qualifiers
                                                                                                                                                                    8.58e-67
922.50
94.368
92.318
89.398
                                                                                                                                                                                     Percent Similarity:
Best Local Similarity:
                                                                                                                                      4032 a
                                                  Unknown
                                                            Unknown
                                                                                                                                                             Alignment Scores:
Pred. No.:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       182
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      893
                                                                                                                                                                                                                                                                                                                                 533
                                                                                                                                                                                                                                                                                                                                                                                                                   102
                                                                                                                                                                                                                                                                                                                                                                                                                                                                   773
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    142
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    833
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     162
                                                                                                                                                                                                                                                'n
                                                                                                                                                                                                                                                               413
                                                                                                                                                                                                                                                                                22
                                                                                                                                                                                                                                                                                               473
                                                                                                                                                                                                                                                                                                                42
                                                                                                                                                                                                                                                                                                                                                  62
                                                                                                                                                                                                                                                                                                                                                                   593
                                                                                                                                                                                                                                                                                                                                                                                   82
                                                                                                                                                                                                                                                                                                                                                                                                  653
                                                                                                                                                                                                                                                                                                                                                                                                                                  713
                                                                                                                                                                                                                                                                                                                                                                                                                                                   122
                                                                                                                                                                                                     Query Match:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        RESULT 2
AR145264
LOCUS
DEFINITION
                                                                                                                      source
                 DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
                                                                                                                                     BASE COUNT
ORIGIN
                                                            ORGANISM
                                                                                    AUTHORS
TITLE
JOURNAL
FEATURES
                                                                            REFERENCE
AR179057
                                                                                                                                                                               Score:
         LOCUS
                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                   ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       셤
                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                               셤
                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                               Q
                                                                                                                                                                                                                                                                                                                õ
                                                                                                                                                                                                                                                                                                                                qq
                                                                                                                                                                                                                                                                                                                                                  ö
                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                   õ
                                                                                                                                                                                                                                                                                                                                                                                                g
```

```
08-AUG-2001
                                                                                                                                                                                                                                                                                                                                                                         243
                                                                                                                                                                                                                                                                                                                                                                                                                                                          363
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 423
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      141
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          483
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             161
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  543
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            603
                                                                                                                                                                                                                                                                                                                                  183
                                                                                                                                                                                                                                                                                                                                                                                                                 303
                                                                                                                                                                                                                                                                                                                                                                                                                                      101
                                                                                                                                                                                                                                                                                                                                                                                                                                                                              121
                                                                                                                                                                                                                                                                                61
                                                                                                                                                                                                                                                                                                                                                                                             83
                                                                                                                                                                                                                                                                                                               22 AlaGlnGlnThrArgGlyGluGluGluGlyCysGlnGluThrSerGlnThrGlyArgAspLys 41
                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ProLysGlyProValThrGlnMetTyrThrAsnValAspLysAspLeuValGlyTrpGln
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       162 AlaGlyHisAlaValGlyIlePheArgAlaAlaValCysThrArgGlyValAlaLysAla
                                                                                                                                                                                                                                                                     GlySerValValIleValGlyArgileAsnLeuSerGlyAsp-----ThrAlaTyr
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              PAT
                                                  1 (bases 1 to 1998)
Malcolm, B.A., Taremi, S.Shane., Weber, P.C. and Yao, N.
Single-chain recombinant complexes of hepatitis C virus
protease and NS4A cofactor peptide
Patent: US 621138-A 105 03-APR-2001;
Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               182 ValAspPheIleProValGluSerLeuGluThrThrMetArgSerPro 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          linear
                                                                                                                                                                         1998
170
16
7
3
                                                                                                                                                                         Length:
Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1998 bp DNA Sequence 109 from patent US 6211338. AR145568 AR145268.1 GI:15107135
                                                                                                                                    u
                                                                                                                                                                                                                               Gaps:
                                                                                                                                                                                                                                                 US-09-965-594-16 (1-197) x AR145264 (1-1998)
                                                                                                                      /organism="unknown"
595 c 569 q
GI:15107131
                                                                                                                                                                         9.21e-67
911.50
94.90%
86.73%
88.32%
                                                                                                              .1998
```

```
Unclassified.
                                                                  410 a
                                                                                     Alignment Scores:
Pred. No.:
                                                                                                                                                                                                                                                                   82
                                                                                                                                                                      64
                                                                                                                                                                                                             42
                                                                                                                                                                                                                                        62
                                                                                                                                                                                                                                                                                304
                                                                                                                                                                                                                                                                                                                                     424
                                                                                                                                                        'n
                                                     source
                                                                                                                                                                                                                                                                                                                                                                                                                                                      DEFINITION
ORGANISM
                                                                  BASE COUNT
            REFERENCE
AUTHORS
TITLE
                                                                                                                                                                                                                                                                                                                                                                                                                                   RESULT 5
AR145263
LOCUS
                                        JOURNAL
                                               FEATURES
                                                                        ORIGIN
                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                        ò
                                                                                                                                                                  g
                                                                                                                                                                                  õ
                                                                                                                                                                                             q
                                                                                                                                                                                                             ò
                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                    a
                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                   셤
                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                             셤
                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                       οy
                                                                                                                                                                                                                                                                                                                                                                                                                                                             08-AUG-2001
                                                                                                                                                                                                                          ValThrargHisAlaaspValIleProValArgArgArgGlyAspSerArgGlySerLeu 141
                                                                                                                                                                                                                                                                                                                                                                LeuSerProArgProlleSerTyrLeuLysGlySerSerGlyGlyProLeuLeuCysPro 161
                                                                                                                                                                     ProLysGlyProValThrGlnMetTyrThrAsnValAspLysAspLeuValGlyTrpGln 101
                                                                                                                                                                                                                                                                                                                                                   PAT
                         1 (bases 1 to 1998)
Malcolm.B.A., Taremi,S.Shane., Weber,P.C. and Yao,N.
Single-chain recombinant complexes of hepatitis C virus
protease and NS4A cofactor peptide
protease and S138-A 109 03-APR-2001;
Location/Qualifiers
1.1998
                                                                                                                                                                                                                                                                                                                                                                                                                    linear
                                                                                                         1998
169
17
7
3
                                                                                                         Length:
Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                           Sequence 103 from patent US 6211338. ARI45262. GI:15107129
                                                                                                                                            Gaps:
                                                                                                                                                        US-09-965-594-16 (1-197) x AR145268 (1-1998)
                                                                        /organism-"unknown"
595 c 569 g
                                                                                                          1.64e-66
908.50
94.90%
86.22%
88.03%
   Unknown.
Unknown.
Unclassified.
                                                                                                                       Percent Similarity:
Best Local Similarity:
                                                                                411 a
                                                                                                    Alignment Scores:
Pred. No.:
                                                                                                                                                                                                                                                                                                                                     122
                                                                                                                                                                                                                                                                                                                                                                                                                     182
                                                                                                                                                                                                                                                                                                                                                                             484
                                                                                                                                                                                                                                                      62
                                                                                                                                                                                                                                                                   244
                                                                                                                                                                                                                                                                                82
                                                                                                                                                                                                                                                                                             304
                                                                                                                                                                                                                                                                                                                                                  424
                                                                                                                                                                                                                                                                                                                                                                142
                                                                                                                                                                                                                                                                                                                                                                                                                                               RESULT 4
AR145262
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
                                                                                                                                     Query Match:
DB:
                                                                   source
             ORGANISM
                                                                               BASE COUNT
                                                    JOURNAL
FEATURES
                                AUTHORS
TITLE
                          REFERENCE
KEYWORDS
                                                                                      ORIGIN
                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                           qq
                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                  a
                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                   셤
                                                                                                                                                                                  a
                                                                                                                                                                                                ŏ
                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                     οy
                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                     අ
                                                                                                                                                                      ò
```

```
08-AUG-2001
                                                                                                                                                                                                                                                                       243
                                                                                                                                                                                                                                                                                                                                                                                  303
                                                                                                                                                                                                                                                                                                                                                                                                     123
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          122 ValThrArgHisAlaAspValIleProValArgArgArgGlyAspSerArgGlySerLeu 141
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               181
                                                                                                                                                                                                                                                                                                                  61
                                                                                                                                                                                                                                                                                                                                                             81
                                                                                                                                                                                                                              21
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             AsnGlnValGluGlyGluValGlnIleValSerThrAlaThrGlnThrPheLeuAlaThr
                                                                                                                                                                                                                                                                                                                                                              CyslleAsnGlyValCysTrpThrValTyrHisGlyAlaGlyThrArgThrlleAlaSer
1 (bases 1 to 1998)
Malcolum, B.A., Taremi, S.Shane., Weber, P.C. and Yao, N.
Single-chain recombinant complexes of hepatitis C virus
protease and NS4A cofactor peptide
Patent: US 6211338-A 103 03-APR-2001;
Location/Qualifiers
1.1998
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         182 ValAspPheIleProValGluSerLeuGluThrThrMetArgSerPro 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             651
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   linear
                                                                                                                              Length:
Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          DNA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Sequence 104 from patent US 6211338.
AR145263
                                                                                     424
                                                                                                                                                                                    Gaps:
                                                                                                                                                                                                          US-09-965-594-16 (1-197) x AR145262 (1-1998)
                                                                         /organism="unknown"
596 c 568 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         GI:15107130
                                                                                                                               1.98e-66
907.50
94.39%
86.73%
87.94%
                                                                                                                                                   Percent Similarity:
Best Local Similarity:
Query Match:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         AR145263.1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Unknown
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Unknown
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
```

Unclassified

Unknown

```
Percent Similarity:
Best Local Similarity:
                                                  ď
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Unknown.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Unknown
                                                                           Alignment Scores:
                                                                                                                Query Match:
DB:
                                   source
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              DEFINITION
ACCESSION
                                                                                  Pred. No.:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ORGANISM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      REFERENCE
AUTHORS
TITLE
                                                  BASE COUNT
                  JOURNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     RESULT 7
AR145266
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              VERSION
KEYWORDS
SOURCE
                           FEATURES
                                                           ORIGIN
                                                                                          Score:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      LOCUS
                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                     a
                                                                                                                                                                                                                                                                                                                                                                      9
                                                                                                                                                                                                                                                                                                                                                                                                   a
                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                  a
                                                                                                                                                                                                                                                                                                                                                                                                                                                     å
                                                                                                                                                                                                                                                                                                                                                                                                                                                                     8
                                                                                                                                                                       a
                                                                                                                                                                                        ð
                                                                                                                                                                                                       g
                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                        δ
                                                                                                                                                                                                                                                                        a
                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                      õ
                                                                                                                                                                                                                                                                                                                                                                                    à
                                                                                                                                                        ð
                                                                                                                                                                                                                         õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      PAT 08-AUG-2001
                                                                                                                                                                                                                                                                                                                                    243
                                                                                                                                                                                                                                                                               ProLysGlyProValThrGlnMetTyrThrAsnValAspLysAspLeuValGlyTrpGln 101
                                                                                                                                                                                                                                       21
                                                                                                                                                                                                        AlaGlnGlnThrArgGlyGluGluGlyCysGlnGluThrSerGlnThrGlyArgAspLys 41
                                                                                                                                                                                                                                                                        81
                                                                                                                                                                         GlySerValValIleValGlyArgIleAsnLeuSerGlyAsp-----ThrAlafyr
                                                                                                                                                                                                                                                                        CysIleAsnGlyValCysTrpThrValTyrHisGlyAlaGlyThrArgThrIleAlaSer
1 (bases 1 to 1998)
Malcolm, B.A., Taremi, S.Shane., Weber, P.C. and Yao, N.
Single-chain recombinant complexes of hepatitis C virus NS3
protease and NS4A cofactor peptide
Patent: US 621338-A 104 03-APR-2001;
Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ValAspPheIleProValGluSerLeuGluThrThrMetArgSerPro 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Unclassified.
1 (bases 1 to 651)
Malcolm,B.A., Taremi,S.Shane., Weber,P.C. and Yao,N.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       linear
                                                                                                1998
170
15
8
3
                                                                                                Length:
Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       DNA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Sequence 95 from patent US 6211338.
AR145254
AR145254.1 GI:15107121
                                                                 424
                                                                                                                                         Gaps:
                                                                                                                                                        (1-1998)
                                                         /organism-"unknown"
596 c 568 g
                                                                   φ
                                                                                                                                                        US-09-965-594-16 (1-197) x AR145263
                                                                                               1.98e-66
907.50
94.39%
86.73%
87.94%
                                                                                                             Percent Similarity:
Best Local Similarity:
Query Match:
                                                                  410 a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Jnknown
                                                                                                                                                                                                                                                                                                                                                                                                                                                                      182
                                                                                                                                                                        S
                                                                                                                                                                                                        22
                                                                                                                                                                                                                                                                       62
                                                                                                                                                                                                                                                                                                       82
                                                                                                                                                                                                                                                                                                                     304
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
                                                   sonrce
                                                                                          Alignment S
Pred. No.:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             DEFINITION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    REFERENCE
AUTHORS
                                                                  BASE COUNT
                                 JOURNAL
FEATURES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     RESULT 6
AR145254
LOCUS
        AUTHORS
                  TITLE
                                                                          ORIGIN
                                                                                                          Score:
                                                                                                                                                                                                                                                                                                                                                                                                                                                    a
                                                                                                                                                                                                                                       δ
                                                                                                                                                                                                                                                      a
                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                     à
                                                                                                                                                                                                                                                                                                                                                    윱
                                                                                                                                                                                                                                                                                                                                                                                   ď
                                                                                                                                                                                                                                                                                                                                                                                                      ð
                                                                                                                                                                                                                                                                                                                                                                                                                     ద
                                                                                                                                                                                                                                                                                                                                                                                                                                    ő
                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                       g
                                                                                                                                                                                                        ò
                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                       å
                                                                                                                                                                                                                                                                                                                     a
                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                        á
```

```
PAT 08-AUG-2001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                       363
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           82 ProLysGlyProValThrGlnMetTyrThrAsnValAspLysAspLeuValGlyTrpGln 101
                                                                                                                                                                                                                                                        64 GGTTCTGTTGTTATTGTTGGTAGAATTATTTTATCTGGTAGTGGTAGTATATCACGGCCTAC 123
                                                                                                                                                                                                                                                                                                                    183
                                                                                                                                                                                                                                                                                                                                                                                                         142 LeuSerProArgProIleSerTyrLeuLysGlySerSerGlyGlyProLeuLeuCysPro 161
                                                                                                                                                                                                                                                                                                                                                                                             81
                                                                                                                                                                                                                                                                                            41
                                                                                                                                                                                                                                                                                                                                             61
                                                                                                                                                                                                                                                                                                                                                                                                                                                           GlySerValVallleValGlyArglleAsnLeuSerGlyAsp-----ThrAlaTyr
                                                                                                                                                                                                                                                                                                                                            42 AsnGlnValGluGlyGluValGlnIleValSerThrAlaThrGlnThrPheLeuAlaThr
                                                                                                                                                                                                                                                                                                                                                         62 CyslleAsnGlyValCysTrpThrValTyrHisGlyAlaGlyThrArgThrIleAlaSer
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1 (bases 1 to 1998)
Malachim, B.A., Taremi, S.Shane., Weber, P.C. and Yao, N.
Single-chain recombinant complexes of hepatitis C virus
protease and NS4A cofactor peptide
    hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                linear
                                                                                                                            651
169
16
7
3
                                                                                                                          Length:
Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                NZ
O
    oŧ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Sequence 107 from patent US 6211338. AR145266
Single-chain recombinant complexes protease and NS4A cofactor peptide Patent: US 6211338-A 95 03-APR-2001
                                                                            144
                                                                                                                                                                                         Gaps:
                                                                                                                                                                                                                 (1-651)
                                      Location/Qualifiers
                                                                /organism="unknown"
187 c 200 g
                                                                                                                                                                                                                  US-09-965-594-16 (1-197) x AR145254
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ARI45266.1 GI:15107133
                                                                                                                         1.03e-66
904.50
94.87%
86.67%
87.65%
                                                    .651
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Unclassified
```

source

#O 4		_	<u> </u>	<u> </u>	<u> </u>		<u> </u>				_ 	O K	4 1 0 4 > X 0 K
JOURNAL Patent: US 6211338-A 107 03-APR-2001; FEATURES Location/Qualifiers Source 11998 ASSE COUNT 410 a 596 c 568 9 424 t ORIGIN	Alignment Scores:  Pred. No.:  Score:  Score:  904.50  Matches:  169  Percent Similarity:  94.394  Conservative:  16  Best Local Similarity:  86.224  Mismatches:  80uery Match:  1 Gaps:  1	S-09-965-594-16 (1-197) x AR145266 (1-1998)	5 GlySerValValIleValGlyArglleAsnLeuSerGlyAspThrAlaTyr 21	22 AlaGlnGlnThrArgGlyGluGluGlyCysGlnGluThrSerGlnThrGlyArgAspLys 41 :::	42 AsnGlnValGluGlyGluValGlnIleValSerThrAlaThrGlnThrPheLeuAlaThr 61	62 CyslleAsnGlyValCysTrpThrValTyrHisGlyAlaGlyThrArgThrIleAlaSer 81    :::	82 ProLysGlyProValThrGlnMetTyrThrAsnValAspLysAspLeuValGlyTrpGln 101	102 AlaProGinGlySerArgSerLeuThrProCysThrCysGlySerSerAspLeuTyrLeu 121 	122 ValThrargHisAlaAspVallleProValArgArgArgGlyAspSerArgGlySerLeu 141 	142 LeuSerProArgProlleSerTyrLeuLysGlySerSerGlyGlyProLeuLeuCysPro 161	162 AlaGlyHisAlaValGlyIlePheArgAlaAlaValCysThrArgGlyValAlaLysAla 181 :::	182 ValAspPheIleProValGluSerLeuGluThrThrMetArgSerPro 197 	RESULT 8 AR145267 LOCUE LOCUE LOCUE LOCUE LOCUE REFERENCE ORGANISM AR145267
FE.	Pro	ΩS	9 9	9 9	å å	oy Oy	9 9	o da	ç, g	9 9	O O	Q D	RE SOE

```
08-AUG-2001
                                                                                                243
                                                                                                                             4.1
                                                                                                                                                         61
                                                                                                                           22 AlaGlnGlnThrArgGlyGluGluGlyCysGlnGluThrSerGlnThrGlyArgAspLys
                                                                                                                                                      1 (bases 1 to 1998)
Malcolm, B.A., Taremi, S.Shane., Weber, P.C. and Yao, N.
Single-chain recombinant complexes of hepatitis C virus NS3
protease and NS4A cofactor peptide
Patent: US 621131318-A 102 03-APR-2001;
Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                   PAT
                                                                                                                                                                                                                                                                                                                                                        ValAspPheIleProValGluSerLeuGluThrThrMetArgSerPro 197
                                                                                                                                                                                                                                                                                                                                                                linear
                                 1998
169
16
8
8
                                  Length:
Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                                 AR145261
Sequence 102 from patent US 6211338.
AR145261.
AR145261.1 GI:15107128
       424
                                                                     Gaps:
                                                                                  JS-09-965-594-16 (1-197) x AR145267 (1-1998)
/organism="unknown"
a 596 c 568 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            /organism-"unknown"
                                 3.53e-66
904.50
94.39%
86.22%
87.65%
                                                                                                                                                                                                                                                                                                                                                                                                                                                   Unclassified.
                                              Percent Similarity:
Best Local Similarity:
Query Match:
DB:
                                                                                                                                                                                                                                                                                                                                                                                                                                     Unknown
                                                                                                                                                                                                                                                                                                                                                                                                                                             Unknown
                            Alignment Scores:
                                                                                                                                                                                                                                                                      122
                                                                                                                                                                                                                                                                                   424
                                                                                                                                                                                                                                                                                                                                                        182
                                                                                                                                                                                                                                                                                                                                                                     604
                                                                                                                                                                                                                                                                                                                                                                                   RESULT 9
AR145261
LOCUS
DEFINITION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      source
      BASE COUNT
ORIGIN
                                                                                                                                                                                                                                                                                                                                                                                                                ACCESSION
VERSION
KEYWORDS
                                                                                                                                                                                                                                                                                                                                                                                                                                           ORGANISM
                                   Pred. No.:
                                                                                                                                                                                                                                                                                                                                                                                                                                                         REFERENCE
AUTHORS
TITLE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      JOURNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              FEATURES
                                                                                                                                                                                                                                                                                                                                                                                                                                     SOURCE
                                                                                                            qq
                                                                                                                                                                                              g
                                                                                                                                                                                                                                       ò q
                                                                                                                                                                                                                                                                               요
장
                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                        δ
                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                 δý
                                                                                                                           ò
                                                                                                                                       q
                                                                                                                                                   à
                                                                                                                                                                  <u>අ</u> දි
                                                                                                                                                                                                           S 6
                                                                                                                                                                                                                                                                   õ
```

us-09-965-594-16.rge

```
Alignment Scores
Pred. No.:
                                                                                                                                                                                           262
                                                                                                                                                                              62
                                                                                                                                                                                                           82
                                                                                                                                                                                                                         322
                                                                                                                                                                                                                                                                                     442
                                              Query Match:
DB:
                                                                                                                                                                                                                                                                                                                                                                                              RESULT 11
AF268278
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                MEDLINE
PUBMED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             REFERENCE
AUTHORS
TITLE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 REFERENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         AUTHORS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      JOURNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        TITLE
                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                    a
                                                                                                                                                                                                                                                                                                   à
                                                                                                                                                                                                                                                                                                                a
                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                             q
                                                                                                            õ
                                                                                                                              g
                                                                                                                                                        qq
                                                                                                                                                                            ò
                                                                                                                                                                                        셤
                                                                                                                                                                                                        ó
                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                  ò
                                                                                    o d
                                                                                                                                           õ
                                                                                                                                                                                                                                                                                                                                                                                                                            08-AUG-2001
                                                                                                                                                                                           363
                                                                                                                                                                                                                                                    483
                                                                                                                                                                                                                                                                                                                   161
                                                                                                                                                                                                                                                                                                                                                 181
                                                                                                         64 GGTTCTGTTGTTATTGTTGGTAGAATTATTTTATCTGGTAGTGGTAGTATCACGCCTAC 123
                                                                                                                                                                            243
                                                                                                                                                                                                                         101
                                                                                                                                       24 TCCCAACAGGGGGGGGGCGTACTTGGTTGCATCATCATCACTTACGCTGGGGGGGAGAG 183
                                                                                                                              22 AlaGlnGlnThrArgGlyGluGluGlyCysGlnGluThrSerGlnThrGlyArgAspLys 41
                                                                                                                                                             61
                                                                                                                                                            ProLysGlyProValThrGlnMetTyrThrAsnValAspLysAspLeuValGlyTrpGln
                                                                                                  5 GlySerValValIleValGlyArgIleAsnLeuSerGlyAsp-----ThrAlaTyr
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1 (bases 1 to 2016)
Malcolm, B.A., Taremi, S.Shane., Weber, P.C. and Yao, N.
Single-chain recombinant complexes of hepatitis C virus
protease and NG4A cofactor peptide
Patent: US 6211338-A 110 03-APR-2001;
                                                                                                                                                                                                                                                                                                                                                                                      ValAspPheIleProValGluSerLeuGluThrThrMetArgSerPro 197
                                                                                                                                                                                                                                                                                                                                                                                                                            linear
                              Length:
Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                      Sequence 110 from patent US 6211338.
AR145269.1 GI:15107136
425
                                                                     Gaps:
                                                                                    US-09-965-594-16 (1-197) x AR145261 (1-1998)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              /organism="unknown"
603 c 570 g
 б
567
                             4.28e-66
903.50
93.88%
86.73%
87.55%
U
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Unclassified.
                                            Percent Similarity:
Best Local Similarity:
Query Match:
DB:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ø
ď
                                                                                                                                                                                                                                                                                                                                                                                                                                                               .
Unknown.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Unknown
409
                        Scores:
                                                                                                                                                                                                                         83
                                                                                                                                                                                                                                                                                                                                                                              182
                                                                                                                                                                                                                                                                                                                                                                                             604
                                                                                                                                                                                                                                                                                                                                                                                                                          LOCUS
DEFINITION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   BASE COUNT
ORIGIN
                       Alignment S
Pred. No.:
                                                                                                                                                                                                                                                                                                                                                                                                                                          ACCESSION
VERSION
KEYWORDS
SOURCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ORGANISM
                                                                                                                                                                                                                                                                                                                                                                                                            RESULT 10
AR145269
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               AUTHORS
TITLE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      JOURNAL
FEATURES
BASE COUNT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         REFERENCE
                                       Score:
                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                             a
                                                                                                                셤
                                                                                                                                                         ò
                                                                                                                                                                         g
                                                                                                                                                                                           ð
                                                                                                                                                                                                       g
                                                                                                                                                                                                                         ογ
                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                    Q
                                                                                                                                                                                                                                                                                    à
                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                 å
                                                                                                                                                                                                                                                                                                                                                                              ö
                                                                                                   ò
                                                                                                                                ö
                                                                                                                                              å
                                                                                                                                                                                                                                                                                                                  ò
```

```
Pestivirus type 1
Pestivirus type 1
Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               VRL 12-JUL-2000
                                                                                                       161
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  561
                                                                                                                                                                              201
                                                                                                                                                                                                                            261
                                                                                                                                                                                                                                                 101
                                                                                                                                                                                                                                                                                                                         381
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          162 AlaGlyHisAlaValGlyIlePheArgAlaAlaValCysThrArgGlyValAlaLysAla 181
                                                                                                                                                       41
                                                                                                                                                                                                      61
                                                                                                                                                    142 LeuSerProArgProIleSerTyrLeuLysGlySerSerGlyGlyProLeuLeuCysPro
                                                                                                                                                                                                                                                                                                                                                                                                                                                       ProLysGlyProValThrGlnMetTyrThrAsnValAspLysAspLeuValGlyTrpGln
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        pestivitus.

1 (bases 1 to 12734)

Lai,V.C., Zhong,W., Skelton,A., Ingravallo,P., Vassilev,V.,
Donis,R.O., Hong,Z. and Lau,J.Y.
Generation and characterization of a hepatitis C virus NS3
protease-dependent bovine viral diarrhea virus
J. Virol. 74 (14), 6339-6347 (2000)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       linear
Length:
Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               RNA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Pestivirus type 1, complete genome.
AF268278
                                                            Gaps:
                                                                                 US-09-965-594-16 (1-197) x AR145269 (1-2016)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                AF268278.1 GI:9049956
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     2 (bases 1 to 12734)
Lai,V.C.H. and Hong,Z.
Direct Submission
4.32e-66
903.50
93.88%
86.73%
                      Percent Similarity:
Best Local Similarity:
```

JOURNAL FEATURES 5'UTR CDS

```
LITHRAY REPRESENTED TO THE CONTROLL OF THE CON
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ARKLGEDVPVYIYATEDEDLAVDLLGLDWPDPGNQQVVETGKALKQVTGLSSAENALL
VALFGYVGYQALSKRIYPMITDIYIEDQRLEDTHHQYAPPAIKTDGTETELKELAS
GDVEKINGGISDYAAGGLEFVKSQAEKIKIAPLFKENAEAAKGYVQKFIDSLIENKEE
ITRYGLMGTHTALYKSJAARLGHETAFATLULKWLAFGGESVSDHYKQAAVDILVYYYY
MNKPSFPGDSETQQEGRRFVASLFISALATYKTWRYHNLSKVVEPALAYLPYATSA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     LKMFTPTRLESVVILSTTIYKTYLSIRKGKSDGLLGTGISAAMEILSQNPVSVGISVM
LGVGAIAAHNAIESSEOKRTLLMKVFVKNFLDOAATDELVKENPEKIIMALFEAVQTI
GNPLRLIYHLYGVYYKGWEAKELSERTAGRNLFTLIMFEAFELLGMDSQGKIRNLSGN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   YILDLIYGLHKQINRGLKKMVLGWAPAPFSCDWTPSDERIRLPTDNYLRVETRCPCGY
EMRAFKNYGGKLTKYEESGPFLCRNRPGRGFYNYRYTKYYDDNLREIKPVAKLEGGYE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              AYTEVNEDVGIIKPVLGERVIPDPVVDINLQPEVQVDISEVGIIIGRETLMTTGVTP
VLEKVEPDASDNQNSVKIGLDEGNYPGPGIQTHTLTEEIHNRDARPFIMILGSRNSIS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          NRAKTARNI NLYTGNDPREIRDLMAAGRMLYVALRDVDPELSEMYDFRGTFLDREALE
ALSLGQPKPKQVTKEAVRNLIEQKKDVEIPNWFASDDPVFLEVALKNDKYYLVGDVGE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        KSNRGHMASAYQLAQGNWEPLGCGVHLGT I PARRYK I HPYEAYLKLKDF I EEEEKKPR
VKDT V I REHNKWI LKK I RPQGNLNTRKMLNPGKLSEQLDREGRKRN I YNHQIGT I MSS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    RGKFWTTLLNGPAFOWCDIGWTGTVSCTSFNADTLATTVVRTYRRSKPFPHRQGCTT

RKKTGEDLHWTLGGRAFVCPEDQLLFKTGGSESCRWGCTGFRESEGTPTV

ENEGGYRLWDSTSCHWEGWALPOGTLKCK IGKTTVQVIAADTKLGBPRCRPYEIIS

EGPVEKTACTFNYTKTLKNKYFEPRDSYFQQYMLKGEYQYWFDLEVTDHHRDYFAEIS

LYTLLRGGRYVLMLLYTWTLSEORAGIGYGSGEVVMADRILTHNYFPLL

LYTLLRESYKKWYLLLYHILWYHPTKSVYTYLLAGGSDYVAADSGGBYLGKTDLCFT

TVVLLVIGLIIARRDPTIVPLVYTIMAALRYTELTHOFGVDIAVAWTITTLLAVSYYTD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    YFRYKKWLOCILSLVSGVFLIRSLIYLGRIEWPEVTIPWRPLTLILLYLISTTIVTR
KVOYAGLIZOCVPILLLYTUMADFTTLILLLYTKTKTYTYFDERSWLGG
IDYRRVDSIYOVDESGEGVYLFPSROKAQGERSILLFLIRATLISCVSSKWOLIYMSY
LTLDFMYYMHRKVIEEISGGTNIISRLVAALIELMWSMEEEESKGLKKFYLLSGRLRN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ERDCATITKNTVQFLKMKKGCAFTYDLTISNLTRLIELVHRNNLEEKEIPTATVTTWL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      /product="polyprotein"
/protein_id="AAR82566.1"
/bc.xref="5g1:9049865.1"
/translation="MELNYPEGSGSVVIVGRIVLSGSGSITACAQOTRGLLGCKITSL
/translation="MELNYPEGSGSVVIVGRIVLSGSGSITACAQOTRGLLGCKITSL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       TVVLAFEEVSQVVKLVLRALRDLTRIWNAATTTAFLICLVKIVRGGMVOGILWLLIIT
GVQGHLDCKPEFSYAIAKDERIGQLGAEGLTTTWKEYSPGMKLEDTWVIAWCEDGKLM
YLORCTRETRYLAILHTRALPTSVVFKKLFDGRRQEDVVEMNDNFEFGLCPCDAKPIV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               HYYKGVTAKIDYSKGKMLLATDKWEVEHGVITRLAKRYTGVGFNGAYLGDEPNHRALV
Submitted (16-MAY-2000) Antiviral Therapy, Schering-Plough Research
Institute, 2015 Galloping Hill Road, Kenilworth, NJ 07033-0539, USA
Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SSGGPLLCPAGHAVGLFRAAVCTRGVAKAVDFIPVENLETTTRSGSGADTEDVVCCSM
SYSDTREEGATKKKTQKPDRLERGKMKIVPKESEKDSKTRPPDATIVVEGVKYQVRKK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              GKTKSKNTODGLYHNKNKPOESRKKLEKALLAMAIIAIVLFOVTMGENITOWNLODNG
TGGIORAMROGGOVNRELIGNEKICTGVDSHLATDIEKELTHGMUDASEKINYTGCR
LORHEWNKHGWGNWYNIEPWILVMNTOQANLTEGOPPRECAYTCRYDRASDLMVYTOA
RDSPTPLTGCKKGKWGNYNESPAGILLARGPCNFEIRASDVLFKEHERISMFODTILYLVDGL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                TNSLEGAROGTAKLTTWLGKOLGILGKKLENKSKTWFGAYAASPYCDVDRKIGYIWYT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        KNCTPACLPKNTKIVGPGKFDTNAEDGKILHEMGGHLSEVLLLSLVVLSDFAPETASV
MYLILHFSIPOSHVDVMDCDKTQLNLTVELTTADVIPGSVWNLGKWVCIRPNWWPYET
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                vkdqakalgatdqtriikevgsrtyamrlsswflqasnkqmsltplfeelllrcppat
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ODLVGWPAPQGSRSLTPCTCGSSDLYLVTRHANVIPVRRRGDSRGSLLSPRPISYLKG
                                                                                                                                                                                                                                                     /organism="Pestivirus type 1"
/mol_type="genomic RNA"
/db_xref="taxon:11099"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            /codon_start-1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                .12508
```

```
PAT 08-AUG-2001
                                                                                                                                                                                                                                                                                                                                                                          772
                                                                                                                                                                                                                                                                                                                                                                                                                                                592
                                                                                                                                                                                                                                                                                              593 TGCATCAATGGGGTATGCTGGACTGTCTACCACGGGGCCGGAACGAGGACCATCGCATCA 652
                                                                                                                                                                                                                                                                                                                                            AlaProGlnGlySerArgSerLeuThrProCysThrCysGlySerSerAspLeuTyrLeu 121
                                                                                                                                                                                                                                                                                                                                                                                                                                        162 AlaGlyHisAlaValGlyIlePheArgAlaAlaValCySThrArgGlyValAlaLySAla 181
                                                                                                                                                                                            532
                                                                                                                                                                                                                                                                                                             ProLysGlyProValThrGlnMetTyrThrAsnValAspLysAspLeuValGlyTrpGln 101
                                                                                                                                                                                                                 41
                                                                                                                                                                                                                                                61
                                                                                                                                                                                                                                                                                81
                                                                                                                                                                                                                                                                                                                                                     GlySerValValIleValGlyArgIleAsnLeuSerGlyAsp-----ThrAlaTyr
                                                                                                                                                                                                                 AlaGlnGlnThrArgGlyGluGluGlyCysGlnGluThrSerGlnThrGlyArgAspLys
                                                                                                                                                                                                                           CyslleAsnGlyValCysTrpThrValTyrHisGlyAlaGlyThrArgThrIleAlaSer
                                                                                                                                                                                                                                                                                                                                                                                                                                                                       182 ValAspPheIleProValGluSerLeuGluThrThrMetArgSer 196
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      997
                                                                               1 others
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      linear
                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Length:
Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      NA
ON
PA
                                                                              ų
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      2802
                                                                                                                                                                   US-09-965-594-16 (1-197) x AF268278 (1-12734)
                                                                                                                                                     Gaps:
                                                                               3293 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            GI:15107125
                                                                                                            3.99e-65
902.50
93.33%
90.77%
87.45%
                                                                              2608 €
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            AR145258.1
                                                                                                                                     Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     AR145258
Sequence 9
AR145258
                                                                                ú
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Unknown.
                                                                               4030
                                                                                                                             Percent Similarity:
                                                                                                      Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                          773
                                                                                                                                                                                                                                                                                                                                                                           122
                                                                                                                                                                                   S
                                                                                                                                                                                                  413
                                                                                                                                                                                                                  22
                                                                                                                                                                                                                                 473
                                                                                                                                                                                                                                                42
                                                                                                                                                                                                                                                                533
                                                                                                                                                                                                                                                                                62
                                                                                                                                                                                                                                                                                                             82
                                                                                                                                                                                                                                                                                                                                             102
                                                                                                                                                                                                                                                                                                                                                            713
                                                                                                                                              Query Match:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
                                                                                                                                     Best Local
                                                                                                               Pred. No.:
                                                                               BASE COUNT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       RESULT 12
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              AR145258
                                                                                      ORIGIN
                                                                                                                      Score:
                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                        a
                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                        셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                       à
                                                                                                                                                                                                                                                                            õ
                                                                                                                                                                                                                                                                                             q
                                                                                                                                                                                                                                                                                                                          d
                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                g
                                                                                                                                                                                                                               q
                                                                                                                                                                                                                                                Ω
                                                                                                                                                                                                                                                            q
                                                                                                                                                                                                                                                                                                             Š
                                                                                                                                                                                   ò
                                                                                                                                                                                                                ο
```

Unknown

6	C. and Xao.N.	TITLE	Single
Single-chain recombinant complexes of protease and NS4A cofactor peptide Patent: US 6211388-89 99 03-APR-2001;	patiti	JOURNAL FEATURES SOUTCE	
1651 /organism="unknown"		BASE COUNT ORIGIN	119
BASE COUNT 120 a 187 C 200 g 144 L ORIGIN		Alignment S	Scores:
Alignment Scores: 1.83e-66 Length: Score: 901.50 Matches: Percent Similarity: 94.87% Conservative: Best Local Similarity: 86.15% Indels: Ouery Match: 6 Gaps:	651 168 17 7 3	ent Si Local / Matc	milarity: Similari h: 594-16 (1
") x AR145258 (1-			5 G1y
5 GlySerValVallleValGlyArglleAsnLeuSerGlyAspThrAlaTyr 	.lyAspThralaTyr 21 	ସ୍ଥ :	64 GGTTC
:::	  CIAGCCITACAGGCGGGACAAG 183	δλ	42 AsnGl
42 ASnGlnValGluGlyGluValGlnIleValSerThrAlaThrGlnThrPheLeuAlaThr	laThrGlnThrPheLeuAlaThr 61	qq	84 AACCA
			62 Cysil
62 CysIleAsnGlyValCysTrpThrValTyrH19GlyAlaGlyThrArgThrIleAlaSer 		<b>q</b>	244 IGCGI
		5 8	
8. FIGUSSELYFICKALINE MELLY INTANDARDARSPLENTALLY INTENDATION OF THE STATE OF THE S			
102 AlabroGlnGlySerArgSerLeuIhrProCysThrCysGlySerSerAspLeuTyrLeu	ysGlySerSerAspLouTyrLeu 121	QQ	364 GCGCC
364 GCGCCCCCGGGGCGCGTTCCTTGACACCATGCACCT	GTGGCAGCTCAGACCTTTACTTG 423	Qy	122 ValTh
122 ValThrArgHisAlaAspVallieProValArgArgArgGlyAspSerArgGlySerLeu	14	<b>q</b> 0	424 GTCAC
424 GICACGAGACAIGCIGACGICAIICCGGIGCGCCGGC	GGGGCGACAGTAGGGGGAGCCTG 483	Qy	142 LeuSe
142 LeuSerProArgProlleSerTyrLeuLysGlySerSerGlyGlyProLeuLeuCysPro	SerGlyGlyProLeuLeuCysPro 161	đ	484 CTCTC
484 CTCTCCCCAGGCCTGTCTCCTACTTGAAGGGCTCTC	screeredricacretecer 543	'n	162 AlaG
162 AlaGlyHisAlaValGlyIlePheArgAlaAlaValCySThrArgGlyValAlaLySAla	SysThrArgGlyValAlaLysAla 181	qo	544 TCGG
544 TCGGGCACGCTGTGGCATCTTCCGGGCTGCCGTAI	recaccessing the control of the cont	č	182 ValAs
182 ValaspphelleProValGluserLeuGluThrThrMetArgSer 	4etArgser 196 	Db RESULT 14	604 GIGG
AR145252 AR145252 651 bp DNA DEFINITION Sequence 93 from patent US 6211338.	linear PAT 08-AUG-2001	LOCUS LOCUS DEFINITION ACCESSION VERSION KEVENDENS	AR14525 Sequend AR14525 AR14525
VERSION AR145252.1 GI:15107119 REPRORDS TO SERVICE TO S		SOURCE ORGANISM	Unknow
ORGANISM DIAKOWN. ORGANISM DIAKOWN. ORGANISM TO A A A A A A A A A A A A A A A A A A		REFERENCE AUTHORS	Malcoln Single

```
PAT 08-AUG-2001
                                                                                                                                                                                                                                                                                                         243
                                                                                                                                                                                                                                                                                                                              31yHisAlaValGlyIlePheArgAlaAlaValCysThrArgGlyValAlaLysAla 181
                                                                                                                                                                                                                    61
                                                                                                                                                                                                                                                 lnGlnThrArgGlyGluGluGlyCysGlnGluThrSerGlnThrGlyArgAspLys 41
                                                                                                                                                                                                                                                                                       erValValIleValGlyArgIleAsnLeuSerGlyAsp-----ThrAlaTyr
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ssified.

Names 1 to 651)

Names, Weber, P.C. and Yao, N.

Names, Meber, P.C. and Yao, N.

e-chain recombinant complexes of hepatitis C virus NS3
Im, B.A., Taremi,S.Shane., Weber,P.C. and Yao,N.
=-chain recombinant complexes of hepatitis C virus ise and NSAA cofactor peptide
:: US 6211338-A 93 03-APR-2001;
Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           linear
                                                                                                                651
169
15
8
3
                                                                                                               Length:
Matches:
Conservative:
Mismatches:
Indels:
Gaps:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         253 651 bp
nce 94 from patent US 6211338.
253
253.1 GI:15107120
                                                                        145
                                                                                                                                                                                      1-197) x AR145252 (1-651)
                                                            /organism="unknown"
188 c 199 g
                                                                                                            2.21e-66
900.50
94.36%
86.67%
87.26%
```

σ

us-09-965-594-16.rge

S4A Cofactor peptide 1138-A 94 03-APR-2001; 1018-A 145 1	FEATURES SOUICE BASE COUNT ORIGIN	Alignment Scores	Score: Score: Percent Similarity: Best Local Similarity Ouery Match: DB:	US-09-965-594-16	Qy 5 GlySerVa	Thralaryr 21 Db 64 GGTTCTGT	41 0b 124	183 Qy 42	61 Db 184	243 Qy 62	81 Db 244 303 Qy 82	101 Db 3	363 Qy 102	Septembrie	141 Db 424	483 . Oy 142	161 Db	543 Qy 162	181 Db 544 603 Qy 182	<b>Q</b>	Search completed: Augu-	PAT 08-AUG-2001							N. virus NS3	irus NS	irus NS
	and NS4A cofactor peptide 18 6211338-A 94 03-APR-2001 Location/Qualifiers 1651	3 a 188 c 199 g 145	2.21e-66 Length: 65 900.50 Matches: 16 94.36% Conservative: 15 86.67% Mismatches: 8	8/.2b% indeis: 6 Gaps:	(1-197) x AR145253																	1998 bp DNA lin 106 from patent US 6211338. 1 GI:15107132	nknown.	nknown.	Unknown. Unclassified.	1998)	o 1998) Taremi,S.Shane., Weber,P.C. and Yao,	o 1998) Taremi, S. Shane., Weber, P.C. and Yao,	o 1998) Taremi,S.Shane., Weber,P.C. and Yao, recombinant complexes of hepatitis C	o 1998) Taremi,S.Shane., Weber,P.C. and Yao, recombinant complexes of hepatitis C	o 1998) Taremi, S.Shane., Weber, P.C. and Yao, recombinant complexes of hepatitis C NSAA cofactor peptide

```
243
                                                                                                                    303
                                                                                                       81
                                                                                          1998
169
15
9
                      Length:
Matches:
Conservative:
Mismatches:
Indels:
Gaps:
          ų
          425
                                             97) x AR145265 (1-1998)
Location/Qualifiers
1. .1998
/organism="unknown"
a 597 c 567 g
                      7.61e-66
900.50
93.88%
86.22%
87.26%
```

gust 31, 2003, 00:46:14 ecs

Run on:

and is derived by analysis of the total score distribution

```
Hepatitis C virus
HCV NS44-NS3 compl
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CDNA encoding hepa
Hepatitis C virus
Hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Combined open read
Hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Combined open read
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Nucleotide sequenc
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Composite hepatiti
                                                                                                                                       Hepatitis C virus
Hepatitis C virus
                                                                                                                                                                                   Hepatitis C virus
Hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Combined open
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Hepatitis; NS3 protease; viral replication; chronic liver disease;
liver failure; liver cancer; mutant; mutein; ds.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Hepatitis C virus NS4A-NS3 fusion protease coding sequence #4.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ALIGNMENTS
                           SUMMARIES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         AAX80363
AAX80347
AAX80348
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                AAX80361
AAX80365
                                                                                                                                                                                                                                                                                                                                                                                                                                                                        AAX80357
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            AAX80351
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     BP.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                AAA73331 standard; DNA; 594
                                                                                              B
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        19-DEC-2000 (first entry)
                                                %
Query
Match Length I
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           651
651
1998
2016
648
648
650
                                                                                                                                                                                                                                                                                                                                                                                                                          612
651
1998
1998
1998
2013
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Hepatitis C virus. Synthetic.
                                                                                                                                     0.001
                                                                                            Score
                                                                                                                                                                                                                                                                                                                                    911.5
908.5
                                                                                                                                                                                                                                                                                                                                                                                                                          904.5
904.5
904.5
904.5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    903.5
901.5
900.5
900.5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                897.5
897.5
897.5
896.5
894
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        893.5
893.5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                900
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            AAA73331;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    RESULT 1
AAA73333
XX
XX
AC AAA7
XX
DT 19-D
DE HEPA
XX
XX
HEPA
                                                                                                                                                                                                                                                                                                                                                                                                                                              Result
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Command line parameters:
-MODEL-frame+pr.model -DEV-xlp
-MODEL-frame+pr.model -DEV-xlp
-MODEL-frame+pr.model -DEV-xlp
-MODEL-frame+pr.model -DEV-xlp
-DB-N.Geneseq_197un03 -OFMT-fastap -SUFFIX-rng -MINMATCH-0.1 -LOOPCL-0
-LOOPEXT-0 -UNITS-Bits -STRT-1 -LATELX-blosum62 -TRANS-human40.cdi
-LIST-45 -DOCALIGN-200 -THR_SCORE-pct -THR_MAX-100 -THR_MIN-0 -ALIGN-15
-MODE-LOCAL -OUTPHT-ptc -NORM-ext -HEAPSIZE-500 -MINLEN-0 -ALIGN-15
-WODE-LOCAL -OUTPHT-ptc -NORM-ext -HEAPSIZE-500 -MINLEN-0 -NAXLEN-200000000
-USER-0SSS94_GGGN_1_11412_grupat_29082003_151918_28302 -NCPD-6 -ICPU-3
-NO.MMAAP -LARGEQUERY -NGG SCORES-0 -WAIT -DSPBLOCK-100 -LONGIGG
-DEV_ITHEOUT-120 -WARN_ITHEOUT-30 -THREADS-1 -XGAPOP-10 -XGAPEXT-0.5 -FGAPOP-6
                                                                                                                                                        (without alignments)
2906.924 Million cell updates/sec
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           N Geneseq_19Jun03:*

| SIDSI/gcgdata/geneseq/geneseqn-embl/Na1980.DAT:*
| SIDSI/gcgdata/geneseq-geneseqn-embl/Na1981.DAT:*
| SIDSI/gcgdata/geneseq-geneseqn-embl/Na1982.DAT:*
| SIDSI/gcgdata/geneseq-geneseqn-embl/Na1983.DAT:*
| SIDSI/gcgdata/geneseq-geneseqn-embl/Na1983.DAT:*
| SIDSI/gcgdata/geneseq-geneseqn-embl/Na1983.DAT:*
| SIDSI/gcgdata/geneseq-geneseqn-embl/Na1985.DAT:*
| SIDSI/gcgdata/geneseq-geneseqn-embl/Na1980.DAT:*
| SIDSI/gcgdata/geneseq-geneseqn-embl/Na1980.DAT:*
| SIDSI/gcgdata/geneseq-geneseqn-embl/Na1980.DAT:*
| SIDSI/gcgdata/geneseq-geneseqn-embl/Na1990.DAT:*
| SIDSI/gcgdata/geneseq-geneseqn-embl/Na2001a.DAT:*
| SIDSI/gcgdata/geneseq-geneseqn-embl/Na2001a.DAT:*
| SIDSI/gcgdata/geneseq-geneseqn-embl/Na2001a.DAT:*
| SIDSI/gcgdata/geneseq-geneseqn-embl/Na2001a.DAT:*
| SIDSI/gcgdata/geneseq-geneseqn-embl/Na2001a.DAT:*
| SIDSI/gcgdata/geneseq-geneseqn-embl/Na2001a.DAT:*
                                                                                                                                                                                                                                                               ......VAKAVDFIPVESLETTMRSP 19
                                                                                                                              August 30, 2003, 19:13:57; Search time 182.939 Seconds
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      /SIDS1/gcgdata/geneseg/genesegn-embl/NA2002.DAT:*/SIDS1/gcgdata/geneseg/genesegn-embl/NA2003.DAT:*
GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

    nucleic search, using frame_plus_p2n model

                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Total number of hits satisfying chosen parameters:
                                                                                                                                                                                                                                                                                                                                                                                                                                          2552756 seqs, 1349719017 residues
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries
                                                                                                                                                                                                                                           1032
1 MKKKGSVVIVGRINLSGDTA.
                                                                                                                                                                                                                                                                                                                                Xgapext
Ygapext
Fgapext
Delext
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Minimum DB seq length: 0
Maximum DB seq length: 2000000000
                                                                                                                                                                                                                      US-09-965-594-16
                                                                                                                                                                                                                                                                                                         BLOSUM62
Xgapop 10.0 , X
Ygapop 10.0 , Y
Fgapop 6.0 , F
Delop 6.0 , D
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Title:
Perfect score:
Sequence:
                                                                                                                                                                                                                                                                                                             Scoring table:
                                                                                       OM protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Database :
                                                                                                                                                                                                                                                                                                                                                                                                                                            Searched:
```

Q

```
The present sequence is the coding sequence for a mutated version of a fusion protein created using the Hepatitis C virus (HCV) NS3 and NS4A protease enzymes. These proteins are both essential for the replication of the virus, acting to cleave its replicative proteins from the polyprotein produced from the HCV genome. Inhibitors of the two proteins should be effective as antiviral treatments of HCV infection. This is useful as HCV can lead to chronic liver disease such as cirrhosis, liver failure and NS3-NS4A fusion proteins which can be used to identify inhibitors of this type, as well as enabling structural studies of the protease and protease:inhibitor complexes. The protein produced from this sequence contains the alpha-helix0-1 variant.
                                                                                                                                                                                                                                                                                                                 Modified hepatitis C virus (HCV) NS3 protease comprising at least 1 substitution of a hydrophobic alpha-helix 0 amino acid to a hydrophilic amino acid, useful for screening inhibitors that may treat hepatitis C
                "NS4A-NS3 fusion protein #4"
                                                                                                                                                                                                                            Goldfarb V;
                                                                                                                                                                                                                            Zhang Y,
                                                                                                                                                                                                                                                                                                                                                                                                          26; Fig 14; 66pp; English.
                                                                                                                                                                                           (BRIM ) BRISTOL-MYERS SQUIBB CO
                                                                                                                                                                                                                            Wittekind M, Weinheimer S,
                                                                                                                      2000WO-US00345
                                                                                                                                                         99US-0115271
 / rag a
/*tag=
                                                                                                                                                                                                                                                               2000-465976/40.
                                                                                                                                                                                                                                                                              P-PSDB; AAB15222
                                                   WO200040707-A1
                                                                                                                                                         08-JAN-1999;
                                                                                                                        06-JAN-2000;
                                                                                                                                                                                                                                                                                                                                                                                                          claim
```

Sequence 594 BP; 105 A; 187 C; 155 G; 147 T; 0 other;

594 197 0 0 0 Conservative: Mismatches: Indels: Length: Matches: Gaps: 8.48e-88 1032.00 100.00% 100.00% 100.00% Percent Similarity: Best Local Similarity: Alignment Scores:

US-09-965-594-16 (1-197) x AAA73331 (1-594)

ò g

9

0, TyralaglnGlnThrArgGlyGluGluGlyCysGlnGluThrSerGlnThrGlyArgAsp 61

9 g

> 셤 ò g ò å

ð

- ThrCysIleAsnGlyValCysTrpThrValTyrHisGlyAlaGlyThrArgThrIleAla 80 61
- 81
- GlnAlaProGlnGlySerArgSerLeuThrProCysThrCysGlySerSerAspLeuTyr 120 101

480 301 CAGGCTCCGCAGGGTTCCCGTTCCCTGACCCCGTGCACCTGCGGTTCCTCCGACCTGTAC 360 ProAlaGlyHisAlaValGlyIlePheArgAlaAlaValCySThrArgGlyValAlaLyS 180 Modified hepatitis C virus (HCV) NS3 protease comprising at least 1 substitution of a hydrophobic alpha-helix  ${\bf 0}$  amino acid to a hydrophilic amino acid, useful for screening inhibitors that may treat hepatitis  ${\bf C}$ LeuValThrArgHisAlaAspValIleProValArgArgArgGlyAspSerArgGlySer LeuLeuSerProArgProIleSerTyrLeuLysGlySerSerGlyGlyProLeuLeuCys Hepatitis; NS3 protease; viral replication; chronic liver disease; liver failure; liver cancer; mutant; mutein; ds. AlavalAspPhelleProValGluSerLeuGluThrThrMetArgSerPro 197 591 Hepatitis C virus NS4A-NS3 fusion protease coding sequence #3. /product- "NS4A-NS3 fusion protein #3" Goldfarb V; Wittekind M. Weinhelmer S. Zhang Y. Location/Qualifiers 1..594 /\*tag= a g BP. (BRIM ) BRISTOL-MYERS SQUIBB 99US-0115271 06-JAN-2000; 2000WO-US00345 AAA73330 standard; DNA; 594 WPI; 2000-465976/40. P-PSDB; AAB15221. Hepatitis C virus. WO200040707-A1 08-JAN-1999; 19-DEC-2000 13-JUL-2000 Synthetic 141 161 AAA73330; 421 181 121 Key RESULT 2 AAA73330 a g ò õ ò

The present sequence is the coding sequence for a mutated version of a fusion protein created using the Hepatitis C virus (HCV) NS3 and NS4A protease enzymes. These proteins are both essential for the replication of the virus, acting to cleave its replicative proteins from the polyprotein produced from the HCV genome. Inhibitors of the two proteins should be effective as antiviral treatments of HCV infection. This is useful as HCV can lead to chronic liver disease such as cirrhosis, liver failure and liver cancer. The present invention concerns a number of NS3 inhibitors of this type, as well as enabling structural studies of the protease and protease:inhibitor complexes. The protein produced from this sequence contains the alpha-helix0-1 variant.

Claim 26; Fig 13; 66pp; English.

 $\sim$ 

```
The present sequence is the coding sequence for a mutated version of a fusion protein created using the Hepatitis C virus (HCV) NS3 and NS4A protease enzymes. These proteins are both essential for the replication of the virus, acting to cleave its replicative proteins from the polyprotein produced from the HCV genome. Inhibitors of the two proteins should be effective as antiviral treatments of HCV infection. This is useful as HCV can lead to chronic liver disease such as cirrhosis, liver failure and Niver cancer. The present invention concerns a number of NS3 mutants and NS3-NS4A fusion proteins which can be used to identify inhibitors of this type, as well as enabling structural studies of the protease and protease inhibitor complexes. The protein produced from this sequence contains the alpha-helix0-1 variant.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          LeuValThrArgHisAlaAspValIleProValArgArgArgGlyAspSerArgGlySer 140
                                                                                                                                                                 Modified hepatitis C virus (HCV) NS3 protease comprising at least 1 substitution of a hydrophobic alpha-helix 0 amino acid to a hydrophilic amino acid, useful for screening inhibitors that may treat hepatitis C
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1 MetLysLysLysGlySerValValIleValGlyArglleAsnLeuSerGlyAspThrAla
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     21 TyrAlaGlnGlnThrArgGlyGluGluGlyCysGlnGluThrSerGlnThrGlyArgAsp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             41 LysAsnGlnValGluGlyGluValGlnIleValSerThrAlaThrGlnThrPheLeuAla
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Sequence 594 BP; 105 A; 189 C; 153 G; 147 T; 0 other;
                                                                                                                                                                                                                                                                                                                                                                                                               594
0 0 0
0 0 0
                                                                                                                                                                                                                                                                                                                                                                                                               Length:
Matches:
Conservative:
                                                                                                                Goldfarb
                                                                                                                                                                                                                                                                                                                                                                                                                                              Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Gaps:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        US-09-965-594-16 (1-197) x AAA73332 (1-594)
                                                                                                                2hang Y,
                                                                                             ဗ
                                                                                                                                                                                                                       Claim 26; Fig 15; 66pp; English
                                                                                           (BRIM ) BRISTOL-MYERS SQUIBB
                                                                                                                                                                                                                                                                                                                                                                                                            5.51e-85
1002.00
98.48%
98.48%
                                                                                                                Weinheimer S,
                                                  06-JAN-2000; 2000WO-US00345
                                                                      99US-0115271
                                                                                                                                   WPI; 2000-465976/40.
P-PSDB; AAB15223.
                                                                                                                                                                                                                                                                                                                                                                                                                                   Percent Similarity:
Best Local Similarity:
        W0200040707-A1
                                                                                                               Wittekind M,
                                                                      08-JAN-1999;
                                                                                                                                                                                                                                                                                                                                                                                                    Alignment Scores:
                            13-JUL-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      121
                                                                                                                                                                                                                                                                                                                                                                                                                                                        Query Match:
                                                                                                                                                                                                                                                                                                                                                                                                                  ..
₽
                                                                                                                                                                                                                                                                                                                                                                                                                Pred.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               엄
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    à
ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                               140
                                                                                                                                                                                                                                                                                                                                                                                                   160
                                                                                                                                                                                                                                                                                                                                                                                                                                             540
                                                                                                                                                                                                                                                                                  240
                                                                                                                                                                                                                                                                                                       100
                                                                                                                                                                                                                                                                                                                            300
                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ProAlaGlyHisAlaValGlyIlePheArgAlaAlaValCysThrArgGlyValAlaLys 180
                                                                                                                                                                                       180
                                                                                                                                                                                                                     9
                                                                                                                                                                                                                                                               80
                                                                                                                                                        9
                                                                                                                                                                             40
                                                                                                                                  ThrCyslleAsnGlyValCysTrpThrValTyrHisGlyAlaGlyThrArgThrIleAla
                                                                                                                                                                                                                                                                                                                                                                                                                                   LeuLeuSerProArgProIleSerTyrLeuLysGlySerSerGlyGlyProLeuLeuCys
                                                                                                                                                                             TyrAlaGlnGlnThrArgGlyGluGluGlyCysGlnGluThrSerGlnThrGlyArgAsp
                                                                                                                                                                                                                     LysAsnGlnValGluGlyGluValGlnIleValSerThrAlaThrGlnThrPheLeuAla
                                                                                                                                                                                                                                 LeuValThrArgHisAlaAspValIleProValArgArgArgGlyAspSerArgGlySer
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Hepatitis; NS3 protease; viral replication; chronic liver disease;
liver failure; liver cancer; mutant; mutein; ds.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Hepatitis C virus NS4A-NS3 fusion protease coding sequence #5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    /product= "NS4A-NS3 fusion protein #5"
        other;
                                       Length:
Matches:
Conservative:
Mismatches:
Indels:
        0
        Ë
       G; 149
                                                                                           Gaps:
                                                                                                                US-09-965-594-16 (1-197) x AAA73330 (1-594)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Location/Qualifiers
        C; 156
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ВP
                                    3.33e-86
1015.00
98.48%
98.48%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  DNA; 594
        A; 186
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            19-DEC-2000 (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Ø
       103
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  AAA73332 standard;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Hepatitis C virus
       BP;
                                                        Percent Similarity:
Best Local Similarity:
Query Match:
       Sequence 594
                            Alignment Scores
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Synthetic
                                                                                                                                                                                                                     41
                                                                                                                                                                             21
                                                                                                                                                                                                                                         121
                                                                                                                                                                                                                                                              61
                                                                                                                                                                                                                                                                                  181
                                                                                                                                                                                                                                                                                                       81
                                                                                                                                                                                                                                                                                                                            241
                                                                                                                                                                                                                                                                                                                                                101
                                                                                                                                                                                                                                                                                                                                                                     301
                                                                                                                                                                                                                                                                                                                                                                                         121
                                                                                                                                                                                                                                                                                                                                                                                                                                  141
                                                                                                                                                                                                                                                                                                                                                                                                                                                       421
                                                                                                                                                                                                                                                                                                                                                                                                                                                                            191
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         541
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     AAA73332;
                                       Pred. No.:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        AAA73332
                                                  Score:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                 à
                                                                                                                                                                                                                                                                                                                                                                                                                                                     a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               셤
                                                                                                                                                                            ò
                                                                                                                                                                                               qq
                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                         8
                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                               8
                                                                                                                                                                                                                                                                                                       ö
                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                         ö
                                                                                                                                                                                                                                                                                                                                                                                                           a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
```

180

80

9

50 9 0

AAA73333

셤

ö

셤 ò

```
US-09-965-594-16 (1-197) x AAA73333 (1-594)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Location/Qualifiers
1..588
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ВР
                    7.34e-84
990.00
97.97%
96.95%
95.93%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         AAA73329 standard; DNA; 588
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Hepatitis C virus.
Synthetic.
                                                                         Similarity:
                                                         Percent Similarity
       Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               19-DEC-2000
                                                                                                                                                                                                                                                                                                                                                                                           61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                 81
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         361
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            141
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               161
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  481
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    101
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            AAA73329
                                                                                          Query Match:
                                                                         Best Local
                        Pred. No.:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Key
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      RESULT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                đ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           οy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ά
                                                                                                                                                                             ò
                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                    ö
                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                      å
                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         The present sequence is the coding sequence for a mutated version of a fusion protein created using the Hepatitis C virus (HCV) NS3 and NS4A protease enzymes. These proteins are both essential for the replication of the virus, acting to cleave its replicative proteins from the polyprotein produced from the HCV genome. Inhibitors of the two proteins should be effective as antiviral treatments of HCV infection. This is useful as HCV can lead to chronic liver disease such as cirrhosis, liver failure and NS3-NS4A fusion proteins which can be used to identify inhibitors of this type, as well as enabling structural studies of the protease and protease inhibitor complexes. The protein produced from this sequence contains the alpha-helix0-7 variant.
                                                                                                                                            Modified hepatitis C virus (HCV) NS3 protease comprising at least 1 substitution of a hydrophobic alpha-helix 0 amino acid to a hydrophilic amino acid, useful for screening inhibitors that may treat hepatitis C
                                                     LeuLeuSerProArgProIleSerTyrLeuLysG1ySerSerG1yG1yProLeuLeuCys
                                                                                                                           ProAlaGlyHisAlaValGly1lePheArgAlaAlaValCysThrArgGlyValAlaLys
421 CIGCIGICCCCCCCCCCCTCCTACCIGAAAGGIICCICCCGGIGGICCCCTGTGC
                                                                                                                                                                                                                                                                                                                                                                                                                                             Hepatítis; NS3 protease; viral replication; chronic liver disease;
liver failure; liver cancer; mutant; mutein; ds.
                                                                                                                                                                                          Hepatitis C virus NS4A-NS3 fusion protease coding sequence #6.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     /*tag= a
/product= "NS4A-NS3 fusion protein #6"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Sequence 594 BP; 104 A; 191 C; 152 G; 147 T; 0 other;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Goldfarb
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Zhang Y,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Location/Qualifiers
1..594
/*tag= a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ဗ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Claim 26; Fig 16; 66pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                (BRIM ) BRISTOL-MYERS SQUIBB
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            06-JAN-2000; 2000WO-US00345
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               99US-0115271
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Wittekind M, Weinheimer S,
                                                                                                                                                                                                                                                                                                    AAA73333 standard; DNA; 594
                                                                                                                                                                                                                                                                                                                                                                          (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 WPI; 2000-465976/40.
P-PSDB; AAB15224.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Hepatitis C virus.
Synthetic.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      W0200040707-A1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               08-JAN-1999;
                                                                                                                                                                                                                                                                                                                                                                        19-DEC-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         13-JUL-2000
                    361
                                                                                                                          161
                                                                                                                                                           481
                                                                                                                                                                                                                                                                                                                                       AAA73333;
```

```
240
                                                                                                                                                                                                                                                                                        100
                                                                                                                                                                                                                                                                                                                                    120
                                                                                                                                                                                                                                                                                                                                                          360
                                                                                                                                                                                                                                                                                                                                                                                 140
                                                                                                                                                                                                                                                                                                                                                                                                     420
                                                                                                                                                                                                                                                                                                                                                                                                                            160
                                                                                                                                                                                                                     180
                                                                                                                                                                                                                                                                                                             300
                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ProAlaGlyHisAlaValGlyIlePheArgAlaAlaValCysThrArgGlyValAlaLys 180
                                                                                                                                                                                               9
                                                                                                                                                                                                                                            80
                                                                                                     20
                                                                                                                           9
                                                                                                                                                                                           LeuLeuSerProArgProlleSerTyrLeuLysGlySerSerGlyGlyProLeuLeuCys
                                                                                                   1 MetLysLysLysGlySerValValIleValGlyArglleAsnLeuSerGlyAspThrAla
                                                                                                                                                                                                                                                                                       SerProLysGlyProValThrGlnMetTyrThrAsnValAspLysAspLeuValGlyTrp
                                                                                                                                                                                                                                                                                                  301 CAGGCICCGCAGGGIICCCGIICCCIGACCCCGIGCACCIGCGGIICCTCCGACCIGIAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                  CIGCTGTCCCCCCCCGCGTCCCCTACCTGAAAGGTTCCTCCGGTGGTCCGCTGCTGTGC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CCGGCTGGTCACGCTGTTGGTATCTTCCGTGCTGTTTCCACCCGTGGTGTTGCTAAA
                                                                                                                1 ATGAAAAAAAAAAGGATCGTTGTTATCGTCGGCCGTATCAACCTGTCCGGTGACACGCT
                                                                                                                                                                                                                                          ThrCysIleAsnGlyValCysTrpThrValTyrHisGlyAlaGlyThrArgThrIleAla
                                                                                                                                                                                                                                                                                                                                                                              LeuValThrArgHisAlaAspValIleProValArgArgArgGlyAspSerArgGlySer
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Hepatitis; NS3 protease; viral replication; chronic liver disease;
liver failure; liver cancer; mutant; mutein; ds.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Hepatitis C virus NS4A-NS3 fusion protease coding sequence #2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           /*tag= a
/product= "NS4A-NS3 fusion protein #2"
594
191
0
0
                     Conservative:
Mismatches:
Indels:
 Length:
Matches:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             WO200040707-A1
```

Ŋ

```
The present sequence is the coding sequence for a mutated version of a fusion protein created using the Hepatitis C virus (HCV) NS3 and NS4A protease enzymes. These proteins are both essential for the replication of the virus, acting to cleave its replicative proteins from the PDJ proteins from the HCV genome. Inhibitors of the two proteins should be effective as antiviral treatments of HCV infection. This is useful as HCV can lead to chronic liver disease such as cirrhosis, liver failure and liver cancer. The present invention concerns a number of NS3 matrants and NS3-NS4A fusion proteins which can be used to identify inhibitors of this type, as well as enabling structural studies of the protease and protease: inhibitor complexes. The protein produced from this
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SerProLysGlyProValThrGlnMetTyrThrAsnValAspLysAspLeuValGlyTrp 100
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   GlnAlaProGlnGlySerArgSerLeuThrProCysThrCysGlySerSerAspLeuTyr 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ThrCysIleAsnGlyValCysTrpThrValTyrHisGlyAlaGlyThrArgThrIleAla 80
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1 MetLysLysLysGlySerValValIleValGlyArglleAsnLeuSerGlyAspThrAla 20
                                                                                                                                                                 Modified hepatitis C virus (HCV) NS3 protease comprising at least 1 substitution of a hydrophobic alpha-helix 0 amino acid to a hydrophilic amino acid, useful for screening inhibitors that may treat hepatitis C
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1 ATGAAAAAAAAGGATCCGTTGTTATCGTCGGCCGTATAGTACTGAACGGT----GCT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Sequence 588 BP; 103 A; 180 C; 156 G; 149 T; 0 other;
                                                                                        Goldfarb V;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Length:
Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                            sequence contains the alpha-helix0-1 variant.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Gaps:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (1-588)
                                                                                        Zhang Y,
                                                             8
                                                                                                                                                                                                                                            Claim 26; Fig 12; 66pp; English
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             US-09-965-594-16 (1-197) x AAA73329
                                                          (BRIM ) BRISTOL-MYERS SQUIBB
                                                                                       Wittekind M, Weinheimer S,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      5.06e-83
06-JAN-2000; 2000WO-US00345
                              990S-0115271
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 981.00
98.95%
96.45%
95.06%
                                                                                                                     WPI; 2000-465976/40.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Percent Similarity:
Best Local Similarity:
Query Match:
                                                                                                                                     P-PSDB: AAB15220
                             08-JAN-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       175
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      81
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   101
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             355
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Pred. No.:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Score:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      à
```

```
The present sequence is the coding sequence for a mutated version of a fusion protein created using the Hepatitis C virus (HCV) NS3 and NS4A protease enzymes. These proteins are both essential for the replication of the virus, acting to cleave its replicative proteins from the polyprotein produced from the HCV genome. Inhibitors of the two proteins should be effective as antiviral treatments of HCV infection. This is useful as HCV can lead to chronic liver disease such as cirrhosis, liver failure and NS3-NS4A fusion proteins which can be used to identify inhibitors of this type, as well as enabling structural studies of the protease and protease:inhibitor complexes. The protein produced from this sequence contains the alpha-helix0-7 variant.
                          474
                                                       ProAlaGlyHisAlaValGlyIlePheArgAlaAlaValCysThrArgGlyValAlaLys 180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Modified hepatitis C virus (HCV) NS3 protease comprising at least 1 substitution of a hydrophobic alpha-helix 0 amino acid to a hydrophilic amino acid, useful for screening inhibitors that may treat hepatitis C
Hepatitis; NS3 protease; viral replication; chronic liver disease;
liver failure; liver cancer; mutant; mutein; ds.
                                                                                                           Hepatitis C virus NS4A-NS3 fusion protease coding sequence #7
                                                                                                                                                                                                                                                                                                                                                                                                                                  "NS4A-NS3 fusion protein #7"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Sequence 594 BP; 105 A; 192 C; 151 G; 146 T; 0 other;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 594
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Goldfarb
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Zhang Y,
                                                                                                                                                                                                                                                                                                                                                                                           Socation/Qualiflers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Claim 26; Fig 17; 66pp; English
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          (BRIM ) BRISTOL-MYERS SQUIBE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Wittekind M, Weinheimer S,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 6.36e-83
980.00
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               99US-0115271
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      06-JAN-2000; 2000WO-US00345
                                                                                                                                                                                             AAA73334 standard; DNA; 594
                                                                                                                                                                                                                                                    (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                        Ø
                                                                                                                                                                                                                                                                                                                                                                                                                                    /product-
                                                                                                                                                                                                                                                                                                                                                                                                        1..594
/*tag=
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  WPI; 2000-465976/40.
                                                                                                                                                                                                                                                                                                                                                   Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                P-PSDB; AAB15225.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                WO200040707-A1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Alignment Scores:
Pred. No.:
Score:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 08-JAN-1999;
                                                                                                                                                                                                                                                    19-DEC-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          13-JUL-2000
                                                                                                                                                                                                                                                                                                                                                                  Synthetic
                                                       161
                                                                                                                                                                                                                           AAA73334;
                                                                                                                                                                   RESULT
                                                                                                                                                                                                q
                                                    ò
                                                                                셤
                                                                                                            Š
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          354
```

```
180
                                                                                                                                                                                                 240
                                                                                                                                                                                                                    140
                                                                                                                                                                                                                                                                                                       LeuLeuSerProArgProlleSerTyrLeuLysGlySerSerGlyGlyProLeuLeuCys 160
                                                                                                                                                                                                                                                                                                                                                    CTGCTGTCCCCCCGCGTCCCGATCTCCTACCTGAAAGGTTCCTCCCGGTGGTCCCCTGCTGTGC 480
                                                                                                                                                                                                                                                                                                                                                                      ProAlaGlyHisAlaValGlyIlePheArgAlaAlaValCysThrArgGlyValAlaLys 180
                                                                                                                                                                                                                                                                                                                                                                                         540
                                                                                                                        120
                                                                                                                                                                                                                                                         GlnAlaProGlnGlySerArgSerLeuThrProCysThrCysGlySerSerAspLeuTyr 120
                                                                                    9
                                                                                                      40
                                                                                                                                           9
                                                                                                                                                                                83
                                                                                                                ThrCysIleAsnGlyValCysTrpThrValTyrHisGlyAlaGlyThrArgThrIleAla
                                                                                                                                                                                           LeuValThrArgHisAlaAspValIleProValArgArgArgGlyAspSerArgGlySer
                                                                                                                                                                                                                                                                                                                                                                                MetLysLysLysGlySerValVallleValGlyArglleAsnLeuSerGlyAspThrAla
                                                                            LysAsnGlnValGluGlyGluValGlnIleValSerThrAlaThrGlnThrPheLeuAla
                                                                                                       TyralaGlnGlnThrArgGlyGluGluGlyCysGlnGluThrSerGlnThrGlyArgAsp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           itis; NS3 protease; viral replication: chronic liver disease;
fallure; liver cancer; mutant; mutein; ds.
                                                                                                                                                                                                                                                                                                                                                                                                            AlavalAspPheIleProValGluSerLeuGluThrThrMetArgSerPro 197
                                                                                                                                                                                                                                                                                                                                                                                                                      Hepatitis C virus NS4A-NS3 fusion protease coding sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       /*tag= a
/product= "NS4A-NS3 fusion protein #8"
  00010
  Conservative:
Mismatches:
Indels:
Gaps:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ΗВ
                                                 US-09-965-594-16 (1-197) x AAA73334
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       06-JAN-2000; 2000WO-US00345
                                                                                                                                                                                                                                                                                                                                                                                                                                                                 AAA73335 standard; DNA; 594
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       (first entry)
 97.46%
96.45%
94.96%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1..594
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Hepatitis C virus.
Synthetic.
Percent Similarity:
Best Local Similarity:
Query Match:
DB:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Hepatitis; NS3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   WO200040707-A1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      19-DEC-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      13-JUL-2000
                                                                                   -
                                                                                                                                                                                                   181
                                                                                                                                                                                                                                       241
                                                                                                                                                                                                                                                                                                               361
                                                                                                                                                                                                                                                                                                                                                                                        481
                                                                                                       21
                                                                                                                         19
                                                                                                                                            41
                                                                                                                                                                                61
                                                                                                                                                                                                                     81
                                                                                                                                                                                                                                                                           301
                                                                                                                                                                                                                                                                                            121
                                                                                                                                                                                                                                                                                                                                  141
                                                                                                                                                                                                                                                                                                                                                   421
                                                                                                                                                                                                                                                                                                                                                                       161
                                                                                                                                                                                                                                                                                                                                                                                                           181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   AAA73335;
                                                                                                                                                                                                                                                         101
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      liver
                                                                                                                                                                                                                                                                                                                                                                                                                                                 RESULT 7
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     셤
                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                   셤
                                                                                                                                                                                                                                                                                                                                                                                        a
                                                                                                                                                                                                                                                                                                                                                                                                           ð
                                                                                                                                                                                                                                                                                                                                                                                                                              a
                                                                   ò
                                                                                   셤
                                                                                                       ò
                                                                                                                        a
                                                                                                                                           ò
                                                                                                                                                           셤
                                                                                                                                                                                ò
                                                                                                                                                                                                g
                                                                                                                                                                                                                    õ
                                                                                                                                                                                                                                      a
                                                                                                                                                                                                                                                          à
                                                                                                                                                                                                                                                                          a
                                                                                                                                                                                                                                                                                            õ
                                                                                                                                                                                                                                                                                                                                                                       à
```

```
The present sequence is the coding sequence for a mutated version of a fusion protein created using the Hepatitis C virus (HCV) NS3 and NS4A protease enzymes. These proteins are both essential for the replication of the virus, acting to cleave its replicative proteins from the polyprotein produced from the HCV genome. Inhibitors of the two proteins should be effective as antiviral treatments of HCV infection. This is useful as HCV can lead to chronic liver disease such as cirrhosis, liver failure and liver cancer. The present invention concerns a number of NS3 mutants and NS3-NS4A fusion proteins which can be used to identify inhibitors of this type, as well as enabling structural studies of the protease and protease:inhibitor complexes. The protein produced from this
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 420
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            40
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                80
                                                                                                                                             Modified hepatitis C virus (HCV) NS3 protease comprising at least 1 substitution of a hydrophobic alpha-helix 0 amino acid to a hydrophilic amino acid, useful for screening inhibitors that may treat hepatitis C
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            21 TyralaGlnGlnThrargGlyGluGluGlyCysGlnGluThrSerGlnThrGlyArgAsp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     41 LysAsnGlnValGluGlyGluValGlnIleValSerThrAlaThrGlnThrPheLeuAla
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     61 ThrCyslleAsnGlyValCysTrpThrValTyrHisGlyAlaGlyThrArgThrIleAla
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             181 ACCIGCATCAACGGIGITIGCIGGACCGITIACCACGGIGCIGGIACCGGIACCAICGCI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        GlnAlaProGlnGlySerArgSerLeuThrProCysThrCysGlySerSerAspLeuTyr
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    121 LeuValThrArgHisAlaAspValIleProValArgArgArgGlyAspSerArgGlySer
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   sequence contains the alpha-helix0 wild-type sequence.
                                                                                                                                                                                                                                                                                                                                                                                                                                                  Sequence 594 BP; 98 A; 189 C; 153 G; 154 T; 0 other;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Length:
Matches:
Conservative:
Mismatches:
Indels:
                                                                       Goldfarb
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Gaps:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   US-09-965-594-16 (1-197) x AAA73335 (1-594)
                                                                       Zhang Y,
                                                                                                                                                                                                                          Disclosure; Fig 18; 66pp; English.
                                            ပ္ပ
                                         (BRIM ) BRISTOL-MYERS SQUIBB
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1.51e-82
976.00
95.94%
95.94%
                                                                       Weinheimer S,
            990S-0115271
                                                                                                  2000-465976/40.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Percent Similarity:
Best Local Similarity:
                                                                                                                  P-PSDB; AAB15226
            08-JAN-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Alignment Scores:
Pred. No.:
                                                                       Wittekind M,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        101
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Query Match:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ď
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Op
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            셤
ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               à
```

```
US-09-965-594-16 (1-197) x AAA73328 (1-588)
                                                                                                                                                                                                                                                                                                                                                                                             415
                                                                                                                                                                                                                                                                                                                                                                                                                    161
                                                                                                                                                                                                                                                                                                                                                                                                                                            475
                                                                                                                                                                                                                                                                                                                                                                      141
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ABA95615;
                                                                                                                                                                                                                         81
                                                                                                                                                                         61
                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                      ö
                                                                                                                                                                                                                                                                                                                                                                                            a
                                                                                                                                                                                                                                                                                                                                                                                                                   Š
                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                          Š
                                               a
                                                                          ô
                                                                                               qq
                                                                                                                         ò
                                                                                                                                                 8
                                                                                                                                                                         ò
                                                                                                                                                                                               g
                                                                                                                                                                                                                       õ
                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                  οχ
                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          The present sequence is the coding sequence for a fusion protein created using the Hepatitis C virus (HCV) NS3 and NS4A protease enzymes. These proteins are both essential for the replication of the virus, acting to cleave its replicative proteins from the polyprotein produced from the HCV genome. Inhibitors of the two proteins should be effective as antiviral treatments of HCV infection. This is useful as HCV can lead to Ahranic liver disease such as cirrhosis, liver failure and liver cancer. The present invention concerns a number of NS3 mutants and NS3-NS4A fusion proteins which can be used to identify inhibitors of this type, as well as embling structural studies of the protease and
                Modified hepatitis C virus (HCV) NS3 protease comprising at least 1 substitution of a hydrophobic alpha-helix 0 amino acid to a hydrophilic amino acid, useful for screening inhibitors that may treat hepatitis C
ProAlaGlyHisAlaValClyIlePheArgAlaAlaValCysThrArgGlyValAlaLys
                                                                                                                                                                                                                      Hepatitis; NS3 protease; viral replication; chronic liver disease;
liver failure; liver cancer; ds.
                                                              AlaValAspPheIleProValGluSerLeuGluThrThrMetArgSerPro 197
                                                                                                                                                                                                 Hepatitis C virus NS4A-NS3 fusion protease coding sequence #1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            97 A: 183 C; 153 G; 155 T; 0 other;
                                                                                                                                                                                                                                                                                                                        /*tag= a /product= "NS3-NS4A fusion protein"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Goldfarb V;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Zhang Y,
                                                                                                                                                                                                                                                                                              Location/Qualifiers
1..588
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Disclosure; Fig 10; 66pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                        (BRIM ) BRISTOL-MYERS SQUIBB CO
                                                                                                                           BP.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    protease:inhibitor complexes.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Wittekind M, Weinheimer S,
                                                                                                                           588
                                                                                                                                                                                                                                                                                                                                                                                                                                 99US-0115271
                                                                                                                                                                                                                                                                                                                                                                                                        06-JAN-2000; 2000WO-US00345
                                                                                                                                                                         (first entry)
                                                                                                                         AAA73328 standard; DNA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    WPI; 2000-465976/40.
P-PSDB; AAB15212.
                                                                                                                                                                                                                                                           Hepatitis C virus.
Synthetic.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Sequence 588 BP;
                                                                                                                                                                                                                                                                                                                                                           WO200040707-A1.
                                                                                                                                                                                                                                                                                                                                                                                                                                 08-JAN-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Alignment Scores:
                                                                                                                                                                         19-DEC-2000
                                                                                                                                                                                                                                                                                                                                                                                   13-JUL-2000
 191
                         181
                                                   181
                                                                          541
                                                                                                                                                 AAA73328
                                                                                                  RESULT 8
                                                                                                                            8
                                                    ô
  6
                          a
```

Matches: Conservative: Mismatches: Indels:

2.29e-79 942.00 94.42% 93.91% 91.28%

Percent Similarity: Best Local Similarity: Query Match: DB:

.. 9

Score:

Length:

```
ProAlaGlyHisAlaValGlyIlePheArgAlaAlaValCysThrArgGlyValAlaLys 180
                                                                                                                                                                                                                                                                                                                                            534
                                                                             SerProLysGlyProValThrGlnMetTyrThrAsnValAspLysAspLeuValGlyTrp 100
                                                                                                                                                                        LeuLeuSerProArgProlleSerTyrLeuLysGlySerSerGlyGlyProLeuLeuCys 160
                                                                                                                                                                                                                                                                                            20
                                       40
                                                                                                                     80
                   54
                                                                                                                                                                                                                                                                                                                                   1 MetLysLysLysGlySerValValIleValGlyArglleAsnLeuSerGlyAspThrAla
                                                                                                                     ThrCysIleAsnGlyValCysTrpThrValTyrHisGlyAlaGlyThrArgThrIleAla
                   1 ATGAAAAAAAAAGGTTCCGTTGTTATCGTCGGCCGTATAGTACTGAACGGT-----GCT
                                       TyralaGlnGlnThrArgGlyGluGluGlyCysGlnGluThrSerGlnThrGlyArgAsp
                                                                                                                                                                                                                                                                                                                                                              Nucleic acid construct encoding chimeric Hepatitis C Virus (RCV)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         qs
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Pestivirus; Npro; protease; NS3; screening;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Chimeric - Bovine viral diarrhea virus.
Chimeric - Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Chimeric BVDV/HCV NS3-wt sequence.
                                                                                                                                                                                                                                                                                                                                                                                                                           BP
                                                                                                                                                                                                                                                                                                                                                                                                                          ABA95615 standard; DNA; 12734
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 99DS-0344456
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    99US-0344456
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Lau JYN;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (SCHE ) SCHERING CORP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Lai VCH,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          US6326137-B1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 25-JUN-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    25-JUN-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  21-MAR-2002
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             04-DEC-2001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Hong 2,
```

œ

us-09-965-594-16.rng

```
pestivirus genome where the Npro protease gene is replaced with NS3 protease gene, useful for in vivo screening of compounds which inhibit HCV infection
```

Example 2; Columns 17-28; 20pp; English.

The present invention relates to a nucleic acid construct encoding a colmaric Hepatitis C virus (HCV)-pestivitus genome. The construct comprises a pestivitus genome where a Npro pestivitus protease gene is replaced with a gene encoding a functional HCV NS3 protease. Furthermore, each junction site recognised by the Npro protease is replaced with a for screening compounds that Inhibit HCV in vivo by inhibiting HCV for screening compounds that Inhibit HCV in vivo by inhibiting HCV the present sequence is a chimaric colone of BVDV (bovine viral diarrhea virus)/HCV NS3-wt, which was used to illustrate the present invention.

HCV; hepatitis C virus; single chain recombinant complex; linker; NS3 protease; NS4A cofactor; NS4A-NS3 complex; serine protease; hydrophobic domain; covalent complex; detection; inhibitor; ss.

Hepatitis C virus

Synthetic

WO9928482-A2

10-JUN-1999

ż

Yao

Weber PC,

Taremi SS,

Malcolm BA,

(SCHE ) SCHERING CORP

98US-0094331.

28-JUL-1998; 28-NOV-1997;

98WO-US24528

24-NOV-1998;

complex encoding cDNA SEQ ID NO:105.

HCV NS4A-NS3

(first entry)

07-SEP-1999

AAX80355;

Sequence 12734 BP; 4032 A; 2604 C; 3295 G; 2803 T; 0 other;

Alignment Scores: Pred. No.: Score: Percent Similarity: Best Local Similarity Ouery Match:	Scores:	es: rity: ilarity:	7.1e-76 922.50 94.368 92.318 89.398	Length: Matches: Conservative: Mismatches: Indels: Gaps:	12734 180 4 8 8 1	
-596-60-SN	- 594 -	16 (1-197)	x ABA95615 (1	-12734)		
Qy	2	GlySerVal	ValileValGlyArg	IleAsnLeuSerGly	GlySerValValIleValGlyArglleAsnLeuSerGlyAspThrAlaTyr	21
qg	413	GGTAGTGTT	GTTATTGTTGGTAGA	ATTGTTTTATCTGGT	AGTGGTAGTATCACGGCGTAC	472
ò	22	AlaGlnGln	ThrArgGlyGluGlu	GlyCysGlnGluThrs	AlaGlnGlnThrArgGlyGluGluGluGlyCysGlnGluThrSerGlnThrGlyArgAspLys	41
qq	473	GCCCAGCAG	ACGAGAGGCCTCCTA	GGGTGTAAGATCACC	AGTCTGACTGGCCGGGACAAA	532
Qy	42	AsnGlnVal	GluGlyGluValGln	IleValSerThrAla	AsnGlnValGluGlyGluValGlnIleValSerThrAlaThrGlnThrPheLeuAlaThr	61
qq	533	AACCAAGTG	GAGGGTGAGGTCCAG	ATCGTGTCAACTGCT	ACCCAAACCTTCCTGGCAACG	592
Qy	62	CyslleAsn	GlyValCysTrpThr	ValTyrHisGlyAla	CyslleAsnGlyValCySTrpThrValTyrHisGlyAlaGlyThrArgThrIleAlaSer	81
qa	593	TGCATCAAT	GGGGTATGCTGGACT	GTCTACCACGGGCCC	GGAACGAGGACCATCGCATCA	652
Oy	83	ProLysGly	ProvalThrGlnMet	TyrThrAsnValAsp	ProLysGlyProValThrGlnMetTyrThrAsnValAsplysAspLeuValGlyTrpGln	101
qq	653	CCCAAGGGT	CCTGTCATCCAGATG	TATACCAATGTGGAC	CAAGACCTTGTGGGCTGGCCC	712
Qy	102	AlaProGln	GlySerArgSerLeu	ThrProCysThrCys	AlaProGlnGlySerArgSerLeuThrProCysThrCysGlySerSerAspLeuTyrLeu	121
q	713	GCTCCTCAA	GGTTCCCGCTCATTC	ACACCCTGCACCTGC	GCTCCTCGGACCTTTACCTG	772
οy	122	ValThrArg	HisAlaAspValIle	ProValArgArgArg	ValThrArgHisAlaAspValIleProValArgArgArgGlyAspSerArgGlySerLeu	141
QQ	773	GTTACGAGG	CACGCCGACGTCATT	CCCGTGCGCCGGCGA	GGTGATAGCAGGGGTAGCCTG	832
Οy	142	LeuSerPro	ArgProlleSerTyr	LeuLysGlySerSer	LeuSerProArgProlleSerTyrLeuLysGlySerSerGlyGlyProLeuLeuCysPro	161
qa	833	CTTTCGCCC	CGGCCCATTCCTAC	CTAAAAGGCTCCTCG	GGGGTCCGCTGTTGTGCCCC	892
Qy	162	AlaGlyHis	AlaValGlyIlePhe	eargalaalayalCys	AlaGlyHisAlaValGlyIlePheArgAlaAlaValCysThrArgGlyValAlaLysAla	181
qq	893	GCGGGACAC	GCCGTGGGCCTATTC	AGGCCGCGCGTGTGC	ACCCGTGGAGTGGCCAAGGCG	952
٥y	182		IleProvalGluSer	ValAspPheIleProValGluSerLeuGluThrThrMetArgSer	ArgSer 196	
qq	953	GIGGACITI	PICCTGIGGAGAAC	GIGGACTITATCCCTGIGGAGAACCTAGAGACAACCATGAGATCC	AGATCC 997	

AAX80355 standard; cDNA; 1998

RESULT 10 AAX80355 ID AAX8 XX

```
243
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          64 GGTTCTGTTATTGTTGGTAGAATTATTTATCTGGTAGTGGTAGTATCACGCCTAC 123
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       124 TCCCAACAGACGCGGGGCCTACTTGGTTGCAAGAAGACTAGCCTTACAGGCCGGGACAAG 183
                                                                                                                                                                                                                                                                                                                                                                                                                                                       The present invention describes a covalent hepatitis C virus (HCV) NS4A-NS3 complex comprising a central hydrophobic domain of native HCV NS4A-NS3 complex comprising a central hydrophobic domain, where the hydrophobic domain of native HCV NS4A peptide is tethered by the linker to the amino terminus of the HCV NS4A peptide is tethered by the linker sequence encodes an example of the above complex. The present NS4A-NS3 complexes are useful for structural determination and determination of mode of binding of HCV inhibitors by NMR spectroscopy. They can also be used for detecting inhibitors by the protease activity and the APPase activity of NS3. The covalent NS4A-NS3 complexes are more soluble, stable and active than the non-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CyslleAsnGlyValCysTrpThrValTyrHisGlyAlaGlyThrArgThrIleAlaSer 81
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              42 AsnGlnValGluGlyGluValGlnIleValSerThrAlaThrGlnThrPheLeuAlaThr 61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     22 AlaGlnGlnThrArgGlyGluGluGlvGlyCysGlnGluThrSerGlnThrGlyArgAspLys 41
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             GlySerValValIleValGlyArgIleAsnLeuSerGlyAsp...---ThrAlaTyr
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       covalent protease-peptide complexes previously available.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1998
170
16
7
3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Sequence 1998 BP; 411 A; 595 C; 569 G; 423 T; 0 other;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Length:
Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                   Disclosure; Page 166-169; 211pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                      New hepatitis C virus covalent complexes
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Gaps:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 US-09-965-594-16 (1-197) x AAX80355 (1-1998)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            7.6e-76
911.50
94.90%
86.73%
88.32%
                                                                                                                                                                                                                                                                                                                                                                           WPI; 1999-385385/32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Percent Similarity:
Best Local Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         62
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Ouery Match:
DB:
ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ó
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   a
```

```
411 A; 595 C; 569 G; 423 T; 0 other;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       BP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      AAX80353 standard; cDNA; 1998
                              1.45e-75
908.50
94.90%
86.22%
88.03%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               07-SEP-1999 (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Hepatitis C virus
Synthetic.
  BP;
                                                  Percent Similarity:
Best Local Similarity:
 Sequence 1998
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             W09928482-A2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 10-JUN-1999.
                      Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           182
                                                                                                                                                                                                                                              62
                                                                                                                                                                                                                                                                                     82
                                                                                                                                                                                                                                                                                                                                                                   122
                                                                                                                                                                                                                                                                                                                                                                                        424
                                                                                                                                                                                                                                                                                                                                                                                                            142
                                                                                                                                                                                                                                                                                                                                                                                                                               484
                                                                                                                                                                                                                                                                                                                                                                                                                                                    162
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          AAX80353;
                                                                                                                        'n
                                                                        Query Match:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    RESULT 12
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                      q
                                                                                                                                                                                                                                                                                                                                                                                                                            qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                    à
                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                     q
                                                                                                                                                                                                                                                                                                                             õ
                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                        ò
                                                                                                                                          ద
                                                                                                                                                               õ
                                                                                                                                                                                음
                                                                                                                                                                                                      à
                                                                                                                                                                                                                       ద
                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                    ô
                                                                               161
                                                                                                                                                                          102 AlaProGlnGlySerArgSerLeuThrProCysThrCysGlySerSerAspLeuTyrLeu 121
                                                            423
                                                                                                                                                               AlaGlyHisAlaValGlyIlePheArgAlaAlaValCysThrArgGlyValAlaLysAla 181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               The present invention describes a covalent hepatitis C virus (HCV) NS4A-NS3 complex comprising a central hydrophobic domain of native HCV NS3 beride protease domain, where the hydrophobic domain of native HCV NS3 serine protease domain, where the hydrophobic domain of native HCV NS3 protease domain. The present to the amino terminus of the HCV NS3 protease domain. The present sequence encodes an example of the above complex. The covalent NS4A-NS3 complexes are useful for structural determination and determination of mode of binding of HCV inhibitors by NAR spectroscopy. They can also be used for detecting inhibitors of the protease activity, the helicase activity and the ATPase activity of NS3. The covalent NS4A-NS3 complexes are more soluble, stable and active than the noncovalent protease-peptide complexes previously available.
                                                                                                                        HCV; hepatitis C virus; single chain recombinant complex; linker; NS3 protease; NS4A cofactor; NS4A-NS3 complex; serine protease; hydrophobic domain; covalent complex; detection; inhibitor; ss.
                                                                                                                                                                                                      ValAspPheIleProValGluSerLeuGluThrThrMetArgSerPro 197
                                                                                                                                                                                                                  NS4A-NS3 complex encoding cDNA SEQ ID NO:109.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ż
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Disclosure; Page 179-182; 211pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            New hepatitis C virus covalent complexes
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Weber PC,
                                                                                                                                                                                                                                                                   AAX80359 standard; cDNA; 1998
                                                                                                                                                                                                                                                                                                                                                                                                                                                               98WO-US24528
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 98US-0094331
97US-0067315
                                                                                                                                                                                                                                                                                                            (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Malcolm BA, Taremi SS,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (SCHE ) SCHERING CORP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        WPI; 1999-385385/32
                                                                                                                                                                                                                                                                                                                                                                                           Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  28-JUL-1998;
28-NOV-1997;
                                                                                                                                                                                                                                                                                                                                                                                                                       WO9928482-A2
                                                                                                                                                                                                                                                                                                                                                                                                                                                               24-NOV-1998;
                                                                                                                                                                                                                                                                                                          07-SEP-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                          10-JUN-1999
                                                                                                                                                                                                                                                                                                                                                                                                    Synthetic
                                                                                                                                           484
                                                                                                                                                                162
                                                                                                                                                                                                         182
                                                                                                    424
                                                                                                                         142
                                                                                                                                                                                                                                                                                       AAX80359
  82
                                                             364
                                                                                                                                                                                                                                               RESULT 11
                                                                                                                                                                                                                                                                    a
                                         ò
                                                             g
                                                                                 ŏ
                                                                                                    g
                                                                                                                         ö
                                                                                                                                            ద
                                                                                                                                                                ŏ
                                                                                                                                                                                   쉽
                                                                                                                                                                                                         ó
  õ
```

```
141
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  483
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  161
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               543
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              423
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             AlaGlyHisAlaValGlyIlePheArgAlaAlaValCysThrArgGlyValAlaLysAla 181
                                                                                                                                                        243
                                                                                                                                                                                                                                                                                                                                             ProLysGlyProValThrGlnMetTyrThrAsnValAspLysAspLeuValGlyTrpGln 101
                                                                                                                                                                                                                                                                                                                                                                                                                                                       102 AlaProGlnGlySerArgSerLeuThrProCysThrCysGlySerSerAspLeuTyrLeu 121
                                                                                                                                                                                                                                                                 61
                                                                                                                                                                                                      4,
                                                                                                                                                                                                                                                                                                                             81
                                                                                                                                       2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             LeuSerProArgProlleSerTyrLeuLysGlySerSerGlyGlyProLeuLeuCysProlleuSerTyrLeuLysGlySerSerGlyGlyGlyProLeuLeuCysPro
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   22 AlaGlnGlnThrArgGlyGluGluGlyCyaGlnGluThrSerGlnThrGlyArgAspLys
                                                                                                                                                                                                                                                                                                                             GlySerValValIleValGlyArglleAsnLeuSerGlyAsp-----ThrAlaTyr
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            MCV: hepatitis C virus; single chain recombinant complex; linker;
NS3 protease; NS4A cofactor; NS4A-NS3 complex; serine protease;
hydrophobic domain; covalent complex; detection; inhibitor; ss.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ValAspPhelleProValGluSerLeuGluThrThrMetArgSerPro 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           HCV NS4A-NS3 complex encoding cDNA SEQ ID NO:103
                             Conservative:
Mismatches:
Indels:
Length:
Matches:
                                                                          Gaps:
                                                                                                         US-09-965-594-16 (1-197) x AAX80359 (1-1998)
```

```
64 GGTTCTCTTGTTATTGTTGGTAGAATTATTTTATCTGGTAGTGGTAGTATCACGGCCTAC 123
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 82 ProLysGlyProValThrGlnMetTyrThrAsnValAspLysAspLeuValGlyTrpGln 101
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 102 AlaProGlnGlySerArgSerLeuThrProCysThrCysGlySerSerAspLeuTyrLeu 121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ValThrArgHisAlaAspValIleProValArgArgArgGlyAspSerArgGlySerLeu 141
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                424 GICACGAGACAIGCIGACGICATICCGGIGCGCGGCGGGGGGGCGACAGIAGGGGGGAGCCIG 483
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    142 LeuSerProArgProIleSerTyrLeuLysGlySerSerGlyGlyProLeuLeuCysPro 161
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         162 AlaGlyHisAlaValGlyIlePheArgAlaAlaValCysThrArgGlyValAlaLysAla 181
                                                                                                                                                                                                                                                                The present invention describes a covalent hepatitis C virus (HCV) NS4A-NS3 complex comprising a central hydrophobic domain of native HCV NS4A periodes domain, where the hydrophobic domain of native HCV NS3 berine protease domain, where the hydrophobic domain of native HCV NS4A peptide is tethered by the linker to the amino terminus of the HCV NS3 protease domain. The present sequence encodes an example of the above complex. The covalent NS4A-NS3 complexes are useful for structural determination and determination of mode of binding of HCV inhibitors by NMR spectroscopy. They can also be used for detecting inhibitors by the protease activity, the helicase activity and the ATPase activity of NS3. The covalent NS4A-NS3 complexes are more soluble, stable and active than the noncovalent protease-peptide complexes previously available.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           22 AlaGlnGlnThrArgGlyGluGluGlyCysGlnGluThrSerGlnThrGlyArgAspLys 41
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       GlySerValValIleValGlyArglleAsnLeuSerGlyAsp-----ThrAlaTyr 21
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                AsnGlnValGluGlyGluValGlnIleValSerThrAlaThrGlnThrPheLeuAlaThr 61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CysIleAsnGlyValCysTrpThrValTyrHisGlyAlaGlyThrArgThrIleAlaSer 81
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Sequence 1998 BP; 410 A; 596 C; 568 G; 424 T; 0 other;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Length:
Matches:
Conservative:
Mismatches:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Indels:
Gaps:
                                                                                                                                                                                                                                Disclosure; Page 160-162; 211pp; English
                                                                                                                                                                                                  New hepatitis C virus covalent complexes
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (1-1998)
                                                                                                                            ξ,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     US-09-965-594-16 (1-197) x AAX80353
                                                                                                                              Weber
98WO-US24528
                                   98US-0094331
97US-0067315
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1.8e-75
907.50
94.39%
86.73%
87.94%
                                                                                                                            Taremi SS,
                                                                                          (SCHE ) SCHERING CORP
                                                                                                                                                              WPI; 1999-385385/32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Percent Similarity:
Best Local Similarity:
Query Match:
24-NOV-1998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Alignment Scores:
                                                                                                                          Malcolm BA,
                                   28-JUL-1998;
28-NOV-1997;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            122
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    42
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        62
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       484
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Pred. No.:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Score:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      a
                                                                                          a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                අ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ç
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  õ
```

```
The present invention describes a covalent hepatitis C virus (HCV) NS4A-NS3 complex comprising a central hydrophobic domain of native HCV NS4A-NS3 complex comprising a central hydrophobic domain of native HCV NS3 earlie protease domain, where the hydrophobic domain of native HCV NS4A peptide is tethered by the linker to the amino terminus of the HCV NS4A peptide is tethered by the linker to the amino terminus of the HCV NS3 protease domain. The present NS4A-NS3 complexes are useful for structural determination and determination of mode of binding of HCV inhibitors by NWR spectroscopy. They can also be used for detecting inhibitors by the protease activity, the helicase activity and the ATPase activity of NS3. The covalent NS4A-NS3 complexes are more soluble, stable and active than the noncovalent protease-peptide complexes previously available.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   21
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         22 AlaGlnGlnThrArgGlyGluGluGlyCysGlnGluThrSerGlnThrGlyArgAspLys 41
5 GlySerValVallleValGlyArglleAsnLeuSerGlyAsp-----ThrAlaTyr
                                                                                                                                                                                                                                                HCV; hepatitis C virus; single chain recombinant complex; linker; NS3 protease; NS4A cofactor; NS4A-NS3 complex; serine protease, hydrophobic domain; covalent complex; detection; inhibitor; ss.
                                        1998
170
15
8
8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                568 G; 424 T; 0 other;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Length:
Matches:
Conservative:
Mismatches:
                                                                                                                                                                                                                      HCV NS4A-NS3 complex encoding cDNA SEQ ID NO:104
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ż
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Disclosure; Page 163-166; 211pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             New hepatitis C virus covalent complexes
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    US-09-965-594-16 (1-197) x AAX80354 (1-1998)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   હ
                                                                                                                                  BP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Sequence 1998 BP; 410 A; 596 C;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Weber
                                                                                                                                AAX80354 standard; cDNA; 1998
                                                                                                                                                                                                                                                                                                                                                                                                                98WO-US24528
                                                                                                                                                                                                                                                                                                                                                                                                                                          980S-0094331
970S-0067315
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1.8e-75
907.50
94.39%
86.73%
87.94%
                                                                                                                                                                                          (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Taremi SS,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (SCHE ) SCHERING CORP.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              WPI; 1999-385385/32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Best Local Similarity:
Query Match:
                                                                                                                                                                                                                                                                                                             Hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Percent Similarity
                                                                                                                                                                                                                                                                                                                                                     WO9928482-A2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                          28-JUL-1998;
28-NOV-1997;
                                                                                                                                                                                                                                                                                                                                                                                                                24-NOV-1998;
                                                                                                                                                                                          07-SEP-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Malcolm BA,
                                                                                                                                                                                                                                                                                                                                                                                  10-JUN-1999
                                                                                                                                                                                                                                                                                                                            Synthetic.
                                                                                                   RESULT 13
                                                                                                                 AAX80354
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Score:
                                                                                                                                                            a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         à
                                           à
```

```
121
                                                                                                                                                                                            ValThrArgHisAlaAspValIleProValArgArgArgGlyAspSerArgGlySerLeu 141
                                                                                                               ProLysGlyProValThrGlnMetTyrThrAsnValAspLysAspLeuValGlyTrpGln 101
                                                                                                                                                                                                                                                                           AlaGlyHisAlaValGlyIlePheArgAlaAlaValCySThrArgGlyValAlaLySAla 181
                                 9
                                                                       CysIleAsnGlyValCysTrpThrValTyrHisGlyAlaGlyThrArgThrIleAlaSer 81
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      has
                                                                                                                                                                                                                                                                                                                                                                                                                                                            Hepatitis C; ds; viral prototoxophore; anti-viral; tumour;
virus; infection; antitumour; toxophore; human immunodeficiency virus;
HIV infection; herpes simplex virus; HSV; rhinovirus; NS3 protease.
                                                                                                                                                      AlaProGlnGlySerArgSerLeuThrProCysThrCysGlySerSerAspLeuTyrLeu
                               AsnGlnValGluGlyGluValGlnIleValSerThrAlaThrGlnThrPheLeuAlaThr
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Novel synthetic viral prototoxophore for treating viral infections, toxin molety incorporated into substrate domain specific for viral enzyme, bound and modified by viral enzyme to get converted into
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          This invention relates to a novel synthetic viral prototoxophore
                                                                                                                                                                                                                                                                                                                  ValAspPheIleProValGluSerLeuGluThrThrMetArgSerPro 197
                                                                                                                                                                                                                                                                                                                            Anti-viral synthetic prototoxophore associated DNA sequence.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             H
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Shepard
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Example 1; Page 62; 66pp; English
                                                                                                                                                                                                                                                                                                                                                                               BP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SIC
                                                                                                                                                                                                                                                                                                                                                                               DNA; 612
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                26-APR-2002; 2002WO-USI3223
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    27-APR-2001; 2001US-286893P
                                                                                                                                                                                                                                                                                                                                                                                                                       (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Cathers BE, Neuteboom
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        (NEWB-) NEWBIOTICS INC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                WPI; 2003-167102/16.
                                                                                                                                                                                                                                                                                                                                                                              ABX15706 standard;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        WO200287500-A2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Jnidentified
                                                                                                                                                                                                                                                                                                                                                                                                                      28-MAR-2003
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             07-NOV-2002,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  toxophore
                              42
                                                                                         244
                                                                                                                                                      102
                                                                                                                                                                        364
                                                                                                                                                                                             122
                                                                                                                                                                                                              124
                                                                                                                                                                                                                                    142
                                                                                                                                                                                                                                                                          162
                                                                                                                                                                                                                                                                                             544
                                                                                                                                                                                                                                                                                                                  182
                                                                                                                                                                                                                                                                                                                                    604
                                                                      62
                                                                                                             82
                                                                                                                                 04
                                                                                                                                                                                                                                                      184
                                                                                                                                                                                                                                                                                                                                                                                                  ABX15706;
                                                                                                                                                                                                                                                                                                                                                                 q
                                                                                                                                                                                                                                                                                                                                   g
                                                                                      Q
                                                                                                             ç
                                                                                                                                 a
                                                                                                                                                  õ
                                                                                                                                                                      q
                                                                                                                                                                                            ò
                                                                                                                                                                                                           셤
                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                  ð
                                                 g
                                                                       ð
                              ð
```

comprising a toxin molety operatively incorporated into a substrate domain specific for a viral enzyme. This prototoxophore may be bound and modified by the viral enzyme. This prototoxophore may be bound coffect of an antiviral agent, this method comprises contacting a cell, infected with a virus or is susceptible to infection, with a prototoxophore. The invention further comprises an assay to identify afteriviral agents, comprising confacting an infected cell with a candidate agent and comparing the ability of the agent to inhibit the candidate agent and comparing the ability of the agent to inhibit the comparing the virus in the cell. The prototoxophores of the invention may have virucide or antitumour activity. The prototoxophores of the invention may be useful for reducing or inhibiting viral infectivity, by contacting a cell (e.g. lymphocyte, or expected with a virus or is susceptible to infection with a virus, with a reflective amount of the prototoxophore. The prototoxophore is also useful for ameliorating the severity of a viral infection in a subject, where the virus is selected from human immunodeficiency virus (HIV), herpes simplex virus (HVV), rhinovirus and hepatitis virus, by administering an effective amount of the prototoxophore to the subject. The prototoxophores of the invention are also useful for reading the severity of a virus prototoxophore to the subject. The prototoxophores of the invention are also useful for treating tumours. The present sequence represents an antiviral prototoxophore associated DNA sequence, this sequence is the invention although it is clearly not a protein agence of the invention and the invention and the present of a prototox prototor of the companient of the present of a prototox prototor and prototox prototor associated DNA sequence, this sequence is the invention all prototox brane of the present of a protein of a protein associated by protein in a subject. 439 CIGTCGCCCCGGCCCATTTCCTACTTGAAAGCTCCTCGGGGGGTCCGCTGTTGTGCCCC 498 139 AACCAAGTGGAGGTGAGGTCCAGATTGTGTCAACTGCTGCCCAAACCTTCCTGGCAACG 198 121 141 AlaGlyHisalaValGlyIlePheArgAlaAlaValCysThrArgGlyValAlaLysAla 181 ProLysGlyProValThrGlnMetTyrThrAsnValAspLysAspLeuValGlyTrpGln 101 LeuSerProArgProlleSerTyrLeuLysGlySerSerGlyGlyProLeuLeuCysPro 161 78 19 GGTAGTGTGGTCATTGTGGGTAGGATCATTTTGTCCGGTAGTGGTAGTATCACGGCGTAC 102 AlaProGlnGlySerArgSerLeuThrProCysThrCysGlySerSerAspLeuTyrLeu 122 ValThrArgHisAlaAspValIleProValArgArgArgGlyAspSerArgGlySerLeu 42 AsnGlnValGluGlyGluValGlnIleValSerThrAlaThrGlnThrPheLeuAlaThr the invention although it is clearly not a protein sequence Sequence 612 BP; 120 A; 171 C; 191 G; 130 T; 0 other; 612 178 3 11 11 Length: Matches: Conservative: Mismatches: Indels: US-09-965-594-16 (1-197) x ABX15706 (1-612) 7.89e-76 904.50 92.82% 91.28% 87.65% Percent Similarity: Best Local Similarity: Alignment Scores: 142 162 82 Query Match: DB: ò q ò g ò d à g 8 q õ a ò a à d

121

```
LeuSerProArgProlleSerTyrLeuLysGlySerSerGlyGlyProLeuLeuCysPro 161
124 TCCCAACAGACGCGGGGCCTACTTGGTTGCAAGAAGACTAGCCTTACAGGCCGGGACAAG 183
                                                                                    ValThrArgHisAlaAspValIleProValArgArgArgGlyAspSerArgGlySerLeu
                                                                                                                                                                                                                                                                                  62 CyslleAsnGlyValCysTrpThrValTyrHisGlyAlaGlyThrArgThrIleAlaSer
                                                                                                                                                                                                                                                                                                                                                                   182
                                                                                                                                                                                                                     122
                                                                                                                                                                                                                                            424
                                                                                                                                                                                                                                                                     142
                                                                                                                                                                                                                                                                                           484
                                                                                                                                                                                                                                            셤
                                                                                                                                                                                                                                                                    à
                                                                                                                                                                                                                                                                                        a
                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                           a
                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                         g
                                                                                                                õ
                                                                                                                                         ద
                                                                                                                                                                     ö
                                                                                                                                                                                        g
                                                                                                                                                                                                              δ
                οy
                                        8
                                                                   õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            The present invention describes a covalent hepatitis C virus (HCV) NS4A-NS3 complex comprising a central hydrophobic domain of native HCV NS3 berihe protease domain, where the hydrophobic domain of native HCV NS4A peptide is tethered by the linker to the amino terminus of the HCV NS4A peptide is tethered by the linker to the amino terminus of the HCV NS3 protease domain. The present sequence encodes an example of the above complex. The covalent NS4A-NS3 complexes are useful for structural determination and determination of mode of binding of HCV inhibitors by NMR spectroscopy. They can also be used for detecting inhibitors by the protease activity, the helicase activity and the ATPase activity of NS3. The covalent NS4A-NS3 complexes are more soluble, stable and active than the noncovalent protease-peptide complexes previously available.
 499 GCGGGGCACGCCGTGGGCATATTTAGGGCCGCGGTGTGCACCCGTGGAGTGGCTAAGGCG 558
                                                                                                                                                                                              HCV; hepatitis C virus; single chain recombinant complex; linker;
NS3 protease; NS4A cofactor; NS4A-NS3 complex; serine protease;
hydrophobic domain; covalent complex; detection; inhibitor; ss.
                       651
169
16
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Sequence 651 BP; 120 A; 187 C; 200 G; 144 T; 0 other;
                                                                                                                                                                        HCV NS4A-NS3 complex encoding cDNA SEQ ID NO:95
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Length:
Matches:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Disclosure; Page 147-148; 211pp; English
                                                                                                                                                                                                                                                                                                                                                                                                                                                                   New hepatitis C virus covalent complexes
                                                                                                                                                                                                                                                                                                                                                                                                                  ρ,
                                                                                                  BP
                                                                                                                                                                                                                                                                                                                                                                                                                   Weber
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           8.52c-76
904.50
94.87%
86.67%
87.65%
                                                                                                 AAX80345 standard; cDNA; 651
                                                                                                                                                                                                                                                                                                                                  98WO-US2452B
                                                                                                                                                                                                                                                                                                                                                         98US-0094331.
                                                                                                                                                (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                   Malcolm BA, Taremi SS,
                                                                                                                                                                                                                                                                                                                                                                                             (SCHE ) SCHERING CORP.
                                                                                                                                                                                                                                                                                                                                                                                                                                          WPI; 1999-385385/32
                                                                                                                                                                                                                                                Hepatitis C virus
                                                                                                                                                                                                                                                                                   WO9928482-A2
                                                                                                                                                                                                                                                                                                                                24-NOV-1998;
                                                                                                                                                                                                                                                                                                                                                        28-JUL-1998;
28-NOV-1997;
                                                                                                                                                                                                                                                                                                         10-JUN-1999
                                                                                                                                                                                                                                                          Synthetic.
                         182
                                                                                                                         AAX80345;
                                                                                     AAX80345
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Pred.
                                                                                                   a
 셤
                           ð
```

94 22

a

ô

AlaGinGlnThrArgGlyGluGluGlyCysGlnGluThrSerGlnThrGlyArgAspLys

GlySerValValIleValGlyArgIleAsnLeuSerGlyAsp-----ThrAlaTyr 21

US-09-965-594-16 (1-197) x AAX80345 (1-651)

Conservative: Mismatches:

Best Local Similarity:

Query Match:

Percent Similarity:

Indels:

```
GIGGACTITGIGCCCGIAGAGICCAIGGAAACIACTAIGCGGICT 648
ValAspPheIleProValGluSerLeuGluThrThrMetArgSer 196
                                                                                                                                        Search completed: August 30, 2003, 19:48:03 Job time : 188.939 secs
```

Title: Perfect score:

.. 0

Sequence:

Scoring table:

Searched:

4, Appli 6, Appli 3, Appli 1, Appli 1, Appli 5, Appli 5, Appli 5, Appli

19, Appl 19, Appl 19, Appl 2, Appl 1, Appli 1, Appli 2, Appli

```
Sequence 19, Appl
Sequence 19, Appl
Sequence 2, Appl1
Sequence 1, Appl1
Sequence 12, Appl1
Sequence 17, Appl
Sequence 17, Appl
Sequence 13, Appl
Sequence 13, Appl
Sequence 16, Appl
Sequence 16, Appl
Sequence 16, Appl
Sequence 15, Appl
Sequence 69, Appl
Sequence 69, Appl
Sequence 69, Appl
Sequence 85, Appl
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        APPLICANT: Withekind, Michael
APPLICANT: Withekind, Michael
APPLICANT: Welnhelmer, Steven
APPLICANT: Welnhelmer, Steven
APPLICANT: Zhang, Yaqun
APPLICANT: Coldfarb, Valentina
TITLE OF INVENTION: Modified Forms of Hepatitis C NS3 Protease for
TITLE OF INVENTION: Facilitating Inhibitor Screening and Structural Studies
TITLE OF INVENTION: of Protease:Inhibitor Complexes
FILE REPERBNCE: DB17Sequences
CURRENT APPLICATION NUMBER: 05/09/965,594
FRIOR APPLICATION NUMBER: 60/115,271
PRIOR APPLICATION NUMBER: 60/115,271
PRIOR APPLICATION NUMBER: 60/115,271
PRIOR FILING DATE: 1999-01-08
NUMBER OF SED ID NOS: 2.0
                                    Sequence 19, A Sequence 21, A Sequence 23, B Sequence 25, B Sequence 4, A Sequence 6, A Sequence 1, A Sequence 1, A
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Seguence 1, Aps Seguence 19,
                                                                                                                                                                                                                                                                                                                              Sequence 1,
Sequence 1,
Sequence 5,
Sequence 5,
Sequence 5,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Sequence Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Sequence (
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Seguence
0. 0S-09-965-594-15
0. 0S-09-965-594-19
0. 0S-09-965-594-13
0. 0S-09-965-594-23
0. 0S-09-965-594-23
0. 0S-09-965-594-25
0. 0S-09-965-594-25
0. 0S-09-965-594-25
0. 0S-09-965-594-25
0. 0S-09-965-594-25
0. 0S-09-965-594-25
1. 0S-09-965-93-19
1. 0S-09-917-563-19
1. 0S-09-919-901-15
1. 0S-0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ALIGNMENTS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Sequence 17, Application US/09965594 Patent No. US20020106642Al GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 TYPE: DNA
CRGANISM: Hepatitis C virus
US-09-965-594-17
                                                                                                                                                                                                                                                                   9646
9646
9646
12980
12980
12980
9379
9416
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            836
10803
10803
9416
9416
2061
13910
13910
2064
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   84.0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          US-09-965-594-17
                                                                                                                                                                                                                              904.5
888.5
888.5
                                                                                                                                                                                                                                                                                                                           881.5
881.5
881.5
878.5
878.5
878.8
878.8
863.5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            863
863
860.5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SEQ ID NO 17
             $\begin{aligned}
$\text{PRESTRICT & $\text{PRESTRIC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Command line parameters:
-MODEL-framet-plan, model -DEV-xlp
-QOCQRQ12_1/USFTO_spool/USO9965594/runat_29082003_151920_28367/app_query.fasta_1.2872
-QO-/CQRQ1_1/USFTO_spool/USO9965594/runat_29082003_151920_28367/app_query.fasta_1.2872
-QD-Published_Applications_NA -OFWT-fastap -SUFFTX-rnpb -MINMATCH-0.1
-LOOPCL-0 -LOOPEXT-0 -UNITS-bits -START-1 -END--1 -MATRIX-blosum62
-TRANS-human40.cdl -LIST-45 -DOCALIGN-200 -THR_SCORE-pet -THR_MAX.END-0
-MAXLEN-200000000 -USER-USO9965594_GCGN_1_1_864_@runat_29082003_151920_28367
-NCPUG-6 -LICD-3 -NO_MAAP_LARREQUERY -NEG_SCORES-0 -WAAT -DSBELCCK-100
-LONGLOG -DEV_TIMEOUT-120 -WARN_TIMEOUT-30 -THREDS-1 -XGAPOP-10 -XGAPEXT-0.5
-FGAPOP-6 -FGAPEXT-7 -YGAPOP-10 -YGAPEXT-0.5 -DELOP-6 -DELEXT-7
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Sequence 17, Appl
                                                                                                                                                                                                                       ; Search time 176.482 Seconds
(without alignments)
2560.981 Million cell updates/sec
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.
                                                                                                                                                                                                                                                                                                                                                                                                                     1 MKKKGSVVIVGRINLSGDTA.......vAKAVDFIPVESLETIMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Description
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  | | cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
| cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
| cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
| cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
| cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
| cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
| cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
| cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
| cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
| cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
| cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
| cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
| cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
| cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
| cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
| cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
| cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
/cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*
                                    GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd
                                                                                                                                                               OM protein - nucleic search, using frame_plus_p2n model
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Total number of hits satisfying chosen parameters:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              594 10 US-09-965-594-17
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1533700 seqs, 1147125425 residues
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Published_Applications_NA:*
                                                                                                                                                                                                                       August 30, 2003, 19:26:03
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Xgapop 10.0 , Xgapext
Ygapop 10.0 , Ygapext
Fgapop 6.0 , Fgapext
Delop 6.0 , Delext
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Minimum DB seq length: 0
Maximum DB seq length: 2000000000
                                                                                                                                                                                                                                                                                                                                                         US-09-965-594-16
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Query
Score Match Length DB
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    BLOSUM62
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1032 100.0
                                                                                                                                                                                                                                                                                                                                                                                          1032
```

Database :

₽.

Appl Appli Appli Appli Appli

Alignment Scores: Pred. No.: Score: Percent Similarity: Best Local Similarity: Ouery Watch:	DB:	US-09-965-594-16 (1-	Qy 1 MetLys	Db 1 ATGAAR	Oy 21 TyrAla	Db 61 TACGCT	Qy 41 LysAsn	121	10 .	181	81	241	101	Db 301 CGGCCI OV 121 LeuVal		141	DD 421 CTGCTC	Qy 161 ProAld	Db 481 CCGC	Oy 181 Alava	RESULT 3  US-09-965-594-19  Sequence 19, Appl.  Patent No. US20021  GENERAL INFORMATION WITHER  APPLICANT: Wither  APPLICANT: Wither  APPLICANT: Gold  TITLE OF INVENTION TITLE OF INVEN
594 197 : 0 0 0		leAsnLeuSerGlyAspThrAla 20	TTCAACCTGTCCGGTGACACCGCT 60		AAACCICCCAGACCGGICGIGAC 120	hralathrGlnthrPheLeuAla 60	Acceraccasaccirciser 180							ArgArgGlyAspSerArgGlySer 140 			AlCysThrArgGlyValAlaLys 180	SITIGCACCCGIGGIGITGCIAAA 540	ChrMetArgSerPro	ACCATGCGTTCCCCG 591	C NS3 Protease for eening and Structural Studics lexes
1.82e-110 Length: 1032.00 Matches: 100.00% Conservative 100.00% Mismatches: 100.00% Indels: 10	x US-09-965-594-17 (1-594)	MetLysLysLysGlySerValValIleValGlyArglleAsnLeuSerGlyAspThrAla	AAAGGATCCGTTGTTATCGTCGGCCGTA	TyralaginginThrargglyGluGluGlyCysGlnGluThrSerGlnThrGlyArgasp	CAGACTCGAGGTGAGGAGGGTTGCCAAG	41 LysAsnGlnValGluGlyGluValGlnIleValSerThrAlaThrGlnThrPheLeuAla	GIIGAAGGIGAAGIICAGAICGIIICCA	ThrCysIleAsnGlyValCysTrpThrValTyrHisGlyAlaGlyThrArgThrIleAla	AACGGTGTTTGCTGGACCGTTTACCACG	SerProLysGlyProvalThrGlnMetTyrThrAsnValAspLysAspLeuValGlyTrp 	GGTCCGGTTACCCAGATGTACACCAACG	GlnAlaProGlnGlySerArgSerLeuThrProCysThrCysGlySerSerAspLeuTyr	CAGGGTTCCCGTGACCCCGTGC	LeuValThrArghisAlaAspVallleFroValArgArgArgArgGlyAspSerArgGlySer 	ProAraProlleSerTvrLenLvs61v9		HisAlaValGlyIlePheArgAlaAla	CCGGCTGGTCACGTATCTTCCGTGCTGCTGTTTGCACCCGTGGTGTTGCTAAA	AlaValAspPheIleProValGluSerLeuGluThrThrMetArgSerPro 197	TTCATCCCGGTTGAATCCCTGGAAACC	RESULT 2 US-09-965-594-15 US-09-965-594-15 US-09-965-594-15 US-09-965-594-15 US-09-965-594-15 US-09-965-594-15 US-09-965-594-15 US-09-965-594 US-09-96-96-96-96-96-96-96-96-96-96-96-96-96
Alignment Scores: Pred. No.: Score: Percent Similarity: Best Local Similarity: Ouery Match:	US-09-965-594-16 (1-197)	Oy 1 MetLysLysI	Db 1 ATGAAAAA	Oy 21 TyrAlaGln(	DD 61 TACGCTCAG	Qy 41 LysAsnGln <sup>1</sup>	Db 121 AAAACCAG	61	181	81	Db 241 TCCCCGAAA	101	301	Oy 121 LeuValThr: 	141	421	Oy 161 ProAlaGly	Db 481 CCGCTGGT	Oy 181 AlaValAsp	Db 541 GCTGTTGAC	RESULT 2 US-09-965-594-15 Sequence 15, Application US/09965594 Patent No. US20020106642A1 GENERAL INFORMATION: APPLICANT: Withekind, Michael APPLICANT: Wainheimer, Steven APPLICANT: Chang, Yaqun APPLICANT: Chang, Yaqun TITLE OF INVENTION: Of Proctease:Inh; FILE OF INVENTION: of Proctease:Inh; FILE REFERENCE: DB17Sequences CURRENT FILING DATE: 2001-09-27 PRIOR APPLICATION NUMBER: 0S/0996; CURRENT FILING DATE: 1999-01-08 NUMBER OF SEQ ID NOS: 26 SOFTWARE: Patentin Ver. 2.0 SEQ ID NO 15 ITREE DATE: DATE: THE CONTENTION OF SEQ ID NOS: 26 SEQ ID NO 15 IENCTH: 594

```
Aftitekind, Michael
Weinheimer, Steven
Zhang, Yaqua
Zhang, Yaqua
Goldfarb, Valentina
FENTION: Modified Forms of Hepatitis C NS3 Protease for
FENTION: Modified Forms of Hepatitis C NS3 Protease for
FENTION: Parilitating Inhibitor Screening and Structural Studies
FENTION: Of Protease:Inhibitor Complexes
ACE: DB17Sequences
LICATION NUMBER: US/09/965,594
LICATION NUMBER: 60/115,271
S DATE: 1999-01-08
S DATE: 1999-01-08
S DATE: 2001 NOS: 26
STEATION VOICE: 20
                                                                                                                                                   180
                                                                                                                                                                                     240
                                                                                                                                                                                                    140
                                                                                                                                                                                                                                                                                                   180
                                                                                                                                                                                                                                                                                                                                                  540
                                                                                                                                                                                                                                                                                    420
                                                                                                     9
                                                                                                                                                                      80
                                                                                      9
                                                                                                                                     594
1194
0
0
0
Length:
Matches:
Conservative:
Mismatches:
Indels:
Gaps:
                                                        -197) x US-09-965-594-15 (1-594)
                                                                                                                                                                                                                                                                                                                                                                                                                 lication US/09965594
20106642A1
1.7e-108
1015.00
98.48%
98.48%
98.35%
                         ty:
```

```
181
                                                                                                                                                                                                                                                                                301
                                                                                                                                                                                                                                                                                                           361
                                                                                                                                                                                                                                                  241
                                                                                                                                                                                                                                                                 101
                                                                                                                                                                                                                                                                                                                           141
                                                                                                                                                                                                                                                                                                                                        421
                                                                                                                                                                                                                                     81
                                                                                                                                                                                                                                                                                             121
                                                                                                                                                                                                                                                                                                        셤
                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                    8
                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                   ð
                                                                                                                              a
                                                                                                                                             ç
                                                                                                                                                         g
                                                                                                                                                                       Q
                                                                                                                                                                                       g
                                                                                                                                                                                                    õ
                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                            Oy
Oy
                                                                                                                                                                                                                                                                                                                          Š
                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                   ó
                                                                                                                                                                                                                                                                                                                                                                                                        Sequence 21, Application US/09965594

Patent No. US20020106642A1

GENERAL INFORMATION:
APPLICANT: Witterfind, Michael
APPLICANT: Witterfind, Michael
APPLICANT: Shang, Yaqun
APPLICANT: Coldfarb, Vabentina
APPLICANT: Goldfarb, Vabentina
TITLE OF INVENTION: Facilitating Inhibitor Screening and Structural Studies
TITLE OF INVENTION: Protease:Inhibitor Complexes
TITLE OF INVENTION: US/09/965,594
CURRENT APPLICATION NUMBER: US/09/965,594
CURRENT APPLICATION NUMBER: 60/115,271
PRIOR FILING DATE: 1999-01-08
NUMBER OF SEQ ID NOS: 26
NUMBER OF SEQ ID NOS: 26
                                                                                                                                                                                                                                                                                                          240
                                                                                                                                                                                                                     180
                                                                                                                                               TACGCTCAGCAGACTCGAGGTGAGGGTTGCCAAGAAACCTCCCAGACCGGTCGTGAC 120
                                                                                                               9
                                                                                                                                TyralaGlnGlnThrargGlyGluGluGluGlyCysGlnGluThrSerGlnThrGlyArgAsp 40
                                                                                                                                                             9
                                                                                                                                                                                          80
                                                                                                   LysAsnGlnValGluGlyGluValGlnIleValSerThrAlaThrGlnThrPheLeuAla
                                                                                                                                                                      594
194
0
3
0
                                   Length:
Matches:
Conservative:
Mismatches:
Indels:
                                                                                       US-09-965-594-16 (1-197) x US-09-965-594-19 (1-594)
                                                                         Gaps:
                                  5.45e-107
1002.00
98.48%
98.48%
97.09%
; TYPE: DNA
; ORGANISM: Hepatitis C virus
US-09-965-594-19
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SOFTWARE: Patentin Ver. 2.0
                                          Score:
Percent Similarity:
Best Local Similarity:
                              Alignment Scores:
Pred. No.:
                                                                                                                                                                                                                                                                                                                                                                                                 RESULT 4
US-09-965-594-21
                                                                                                                                                                                                                                                                                                                                                      481
                                                                                                                                                                                                                                                                                                          141
                                                                                                                                  21
                                                                                                                                               19
                                                                                                                                                                                                                                    241
                                                                                                                                                                                                                                                                                             361
                                                                                                                                                                                                                                                                                                                                        161
                                                                                                                                                             41
                                                                                                                                                                           121
                                                                                                                                                                                          61
                                                                                                                                                                                                        181
                                                                                                                                                                                                                      81
                                                                                                                                                                                                                                                   101
                                                                                                                                                                                                                                                                 301
                                                                                                                                                                                                                                                                                121
                                                                                                                                                                                                                                                                                                                          421
                                                                 Query Match:
DB:
                                                                                                     ö
                                                                                                                  g
                                                                                                                                  õ
                                                                                                                                               g
                                                                                                                                                                        a
                                                                                                                                                                                                    g
                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                              ద
                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                       ô
                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                     a
                                                                                                                                                                                                                                                                                                                                                                     ö
                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                          οy
                                                                                                                                                                                        ò
```

```
Studies
                                                                                                                                                                                                                                                                                                                                                                                                                                                      240
                                                                                                                                                                                                                                                                                                                                                       80
                                                                                                                                                                                                      20
                                                                                                                                                                                                                              9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Sequence 13, Application US/09965594

Patent No. US20020106642al

GENERAL INCOMATION:
APPLICANT: Wainhelmer, Steven
APPLICANT: Wainhelmer, Steven
APPLICANT: Wainhelmer, Steven
APPLICANT: Chang, Yaqin
APPLICANT: Chang, Yaqin
APPLICANT: Chang, Yaqin
TITLE OF INVENTION: Pacilitating Inhibitor Screening and Structural Stud
TITLE OF INVENTION: Pacilitating Inhibitor Complexes
TITLE OF INVENTION: Of Protease:Inhibitor Complexes
CURRENT APPLICATION NUMBER: US/09/965,594
CURRENT APPLICATION NUMBER: US/09/115,271
PRIOR APPLICATION NUMBER: 60/115,271
PRIOR FILING DATE: 1999-01-08
                                                                                                                                                                                                                                                                                                                                                                    MetLysLysLysGlySerValValIleValGlyArgIleAsnLeuSerGlyAspThrAla
                                                                                                                                                                                                                    61 ThrCyslleAsnGlyValCysTrpThrValTyrHisGlyAlaGlyThrArgThrlleAla
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       594
191
2
0
0
                                                                                                                  Conservative:
Mismatches:
Indels:
                                                                                                                                                                               US-09-965-594-16 (1-197) x US-09-965-594-21 (1-594)
                                                                                          Length:
                                                                                        1.34e-105
990.00
97.97%
96.95%
95.93%
                                          virus
; SEQ ID NO 21
; LENGTH: 594
; TYPE: DNA
; ORGANISM: Hepatitis C
US-09-965-594-21
                                                                                                               Percent Similarity:
Best Local Similarity:
Query Match:
DB:
                                                                            Alignment Scores:
Pred. No.:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        US-09-965-594-13
```

```
RESULT 7
US-09-965-594-25
                                                                  Alignment Scores
                                                                                                                                                                                                                                             81
                                                                                                                                                                                                                                                           241
                                                                                                                                                                                                                                                                        101
                                                                                                                                                                                                                                                                                     301
                                                                                                                                                                                                                                                                                                   121
                                                                                                                                                                                                                                                                                                                 361
                                                                                                                                                                                                                                                                                                                               141
                                                                                                                                                                                                                                                                                                                                           421
                                                                                                                                                                                                                                                                                                                                                                      481
                                       DNA
                                                                         Pred. No.:
                                                                               Score:
                                                                                                                                                                                                                                                                   õ
                                                                                                                                                                                                                                                                                   a
                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                              셤
                                                                                                                                                                                                                                                                                                                            5
                                                                                                                                                                                                                                                                                                                                           a
                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                   q
                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                a
                                                                                                                                               a
                                                                                                                                                            ŏ
                                                                                                                                                                         셤
                                                                                                                                                                                                    g
                                                                                                                                                                                                                    ð
                                                                                                                                                                                                                            셤
                                                                                                                                                                                                                                             õ
                                                                                                                                                                                                                                                         유
                                                                                                                                     ò
                                                                                                                                                                                    5
                                                                                                                                                                                                                                                                                                                                                                                                         Sequence 23, Application US/09965594

Patent No. US20020106642a1

GENERAL INFORMION:

APPLICANT: Wittekind, Michael

APPLICANT: Wittekind, Michael

APPLICANT: Walnelmer, Steven

APPLICANT: Chang, Yaqun

APPLICANT: Chang, Yaqun

TITLE OF INVENTION: Pacilitating Inhibitor Screening and Structural Studies

TITLE OF INVENTION: Of Protease:Inhibitor Complexes

FILE REFERENCE: DB/7Sequences

CURRENT APPLICATION NUMBER: US/09/965,594
                                                                                                                                                                                                                                160
                                                                                                                                                                                                                                                                                                                                             180
                                                                                                                                                 414
                                                                                                                                                                                                                                                                                                                              474
                                                                                                                                                                                                                                                                                                                                                         534
                                                                                                                                                                                         174
                                                                                                                                                                                                                    234
                                                                                                                                                                                                      80
                                                                                                                        20
                                                                                                                                                                            9
                                                                                                                                    54
                                                                                                                                                                           LeuLeuSerProArgProIleSerTyrLeuLysGlySerSerGlyGlyProLeuLeuCys
                                                                                                                                                                                                                                                                                                                                             ProAlaGlyHisAlaValGlyIlePheArgAlaAlaValCysThrArgGlyValAlaLys
                                                                                                                                                                                                                                                                                                                                                                      588
190
14
12
                                                                         Conservative:
Mismatches:
Indels:
                                                                                                          US-09-965-594-16 (1-197) x US-09-965-594-13 (1-588)
                                                            Length:
Matches:
                                                                                            Gaps:
                                                            1.46e-104
981.00
96.95%
96.45%
95.06%
                                  C virus
NUMBER OF SEQ ID NOS: 26
SOFTWARE: PatentIn Ver.
SEQ ID NO 13
LENGTH: 588
                        ; TYPE: DNA
; ORGANISM: Hepatitis
US-09-965-594-13
                                                                         Percent Similarity:
Best Local Similarity:
Query Match:
                                                                                                                                                                                                                                                                                                                                                                                                RESULT 6
US-09-965-594-23
                                                     Alignment Scores: Pred. No.:
                                                                                                                                                                                          115
                                                                                                                                                                                                                                                                                                                              415
                                                                                                                                                                                                                                                                                                                                             161
                                                                                                                                                                            41
                                                                                                                                                                                                        6.1
                                                                                                                                                                                                                    175
                                                                                                                                                                                                                                 81
                                                                                                                                                                                                                                              235
                                                                                                                                                                                                                                                                                      121
                                                                                                                                                                                                                                                                                                   355
                                                                                                                                                                                                                                                                                                                 141
                                                                                                                                                                                                                                                                                                                                                         475
                                                                    Score:
                                                                                                                                                                                                                                                                                                                              d
                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                         a
                                                                                                                                                                                                                                                                                                                                                                      òγ
                                                                                                                                                                                                                                                                                                                                                                                    合
                                                                                                                                                                                                     ŏ
                                                                                                                                                                                                                    셤
                                                                                                                                                                                                                                 à
                                                                                                                                                                                                                                             a
                                                                                                                                                                                                                                                            à
                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                      õ
                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                        ò
                                                                                                                                    g
                                                                                                                                                                            δ
                                                                                                                                                                                       q
                                                                                                                                                   ò
                                                                                                                                                              g
```

```
Protease for and Structural Studies
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           161 ProAlaGlyHisAlaValGlyIlePheArgAlaAlaValCysThrArgGlyValAlaLys
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                SerProLysGlyProvalThrGlnMetTyrThrAsnValAspLysAspLeuValGlyTrp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           LeuLeuSerProArgProlleSerTyrLeuLysG1ySerSerG1yG1yProLeuLeuCys
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CCGGCTGGTCACGCTGTTGGTATCTTCCGTGCTGCTGTTTCCACCCGTGGTGTTGCTAAA
                                                                                                                                                                                                                                                                                                                        CTGCTGTCCCCCCCCTCCGATCTCCTACCTGAAAGGTTCCTCCGGTGGTCCGCTGCTGTGC
                                                                                                                                                                                                                                                                                                                                                                                       21 TyralaGlnGlnThrArgGlyGluGluGlyCysGlnGluThrSerGlnThrGlyArgAsp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Sequence 25, Application US/09965594

Patent No. US20020106642A1

GENERAL INFORMATION:
GENERAL INFORMATION:
APPLICANT: Withhelmer, Steven

APPLICANT: Welnhelmer, Steven

APPLICANT: Welnhelmer, Steven

APPLICANT: Goldfarb, Valentina

TITLE OF INVENTION: Modified Forms of Hepatitis C NS3 Prot

TITLE OF INVENTION: Facilitating Inhibitor Screening and

TITLE OF INVENTION: of Protease:Inhibitor Complexes

FILE REFERENCE: DB17Sequences
                                                                                                                                                                            Length:
Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                            US-09-965-594-16 (1-197) x US-09-965-594-23 (1-594)
PRIOR APPLICATION NUMBER: 60/115,271
PRIOR FILING DATE: 1999-01-08
NUMBER OF SEQ ID NOS: 26
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 23
LENGTH: 594
                                                                                                                                                                              1.93e-104
980.00
97.46%
96.45%
                                                                                                                 virus
                                                                                                                 U
                                                                                                                 ; ORGANISM: Hepatitis
US-09-965-594-23
                                                                                                                                                                                                     Percent Similarity:
Best Local Similarity:
Query Match:
DB:
```

9

9 40 240 100 300

80

140

420 160 480

180

```
US-09-965-594-4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     355
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               475
                                                                                                                                                                                                                                                                                                                                                                                                                             101
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             161
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     535
                                                                                                                                                                                    Query Match:
DB:
                                                                                                                                               Pred. No.:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                   ŏ
                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                      a
                                                                                                                                                                                                                                                                                                              õ
                                                                                                                                                                                                                                                                                                                              ద
                                                                                                                                                                                                                                                                                                                                                 οy
                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                            οy
                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     APPLICANT: Zhang, Yaqun APPLICANT: Zhang, Yaqun APPLICANT: Goldfarb, Valentina TITLE OF INVENTION: Modified Forms of Hepatitis C NS3 Protease for TITLE OF INVENTION: Facilitating Inhibitor Screening and Structural Studies
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    540
                                                                                                                                                                                                                                                     101 GlnalaProGlnGlyScrArgSerLeuThrProCysThrCysGlySerSerAspLeuTyr 120
                                                                                                                                                                                                                                                                                                                                                                                                                             360
                                                                                                                                                                                                                                                                                                                                                                                                                                               ProAlaGlyHisAlaValGlyIlePheArgAlaAlaValCysThrArgGlyValAlaLys 180
                                                                                                                                                                                                                                                                                                              180
                                                                                                                                                                                                                                                                                                                                                                     SerProLysGlyProValThrGlnMetTyrThrAsnValAspLysAspLeuValGlyTrp 100
                                                                                                                                                                                                                                    9
                                                                                                                                                                                                                                                                                           9
                                                                                                                                                                                                                  20
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CCGGCTCCGCAGGGTTCCCGTTCCCTGACCCCGTGCACCTGCGGTTCCTCCGACCTGTAC
                                                                                                                                                                                                                            1 ATCAAAAAAAAGGATCCGTTGTTATCGTCGCCGTATCAACCTGTCCGGTGACGCCT
                                                                                                                                                                                                                                                                                         LysasnGlnValGluGlyGluValGln1leValSerThralaThrGlnThrPheLeuAla
                                                                                                                                                                                                                                                                                                     MetLysLysLysGlySerValValIleValGlyArgIleAsnLeuSerGlyAspThrAla
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                AlaValAspPhelleProValGluSerLeuGluThrThrMetArgSerPro 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          594
189
0
0
0
                                                                                                                            Length:
Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                                                                                               US-09-965-594-16 (1-197) x US-09-965-594-25 (1-594)
CURRENT APPLICATION NUMBER: US/09/965,594
CURRENT FILING DATE: 2001-09-27
PRIOR APPLICATION UNDER: 60/115,271
PRIOR FILING DATE: 1999-01-08
NUMBER OF SEQ ID NOS: 26
SEQ ID NO 25
LENGTH: 594
                                                                                                                                                                          Gaps:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Sequence 4, Application US/09965594
Patent No. US20020106642A1
GENERAL INFORMATION:
APPLICANT: Wittekind, Michael
APPLICANT: Weinheimer, Steven
                                                                                                                            5.62e-104
976.00
95.94%
95.94%
94.57%
                                                                             ; TYPE: DNA
; ORGANISM: Hepatitis C virus
US-09-965-594-25
                                                                                                                                           Percent Similarity:
Best Local Similarity:
Query Match:
                                                                                                                   Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       RESULT 8
US-09-965-594-4
                                                                                                                                                                                                                                                                                           4.1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             481
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                181
                                                                                                                                                                                                                                                       21
                                                                                                                                                                                                                                                                        61
                                                                                                                                                                                                                                                                                                              121
                                                                                                                                                                                                                                                                                                                                 61
                                                                                                                                                                                                                                                                                                                                                  181
                                                                                                                                                                                                                                                                                                                                                                     81
                                                                                                                                                                                                                                                                                                                                                                                        241
                                                                                                                                                                                                                                                                                                                                                                                                                             301
                                                                                                                                                                                                                                                                                                                                                                                                                                                121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  361
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     141
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        421
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           161
                                                                                                                              Pred. No.:
                                                                                                                                      Score:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                à
                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            à
                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                            셤
                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                Q
                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                    a
                                                                                                                                                                                                                                                                                                                                                                                                        ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                               õ
                                                                                                                                                                                                                  õ
```

```
ProAlaGlyHisAlaValGlyIlePheArgAlaAlaValCysThrArgGlyValAlaLys 180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              295 CCGGCTCCGCAGGTTCCCGTTCCCTGACCCCGTGCACCTGCGGTTCCTCCGACCTGAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            55 TACGCTCAGCAGACTCGAGGTCTGCTGGGTTGCATCATCACCTCCCTGACCGGTCGTGAC 114
                                                                                                                                                                                                                                                                                                                                                                                                                                 21 TyrAlaGloGloThrArgGlyGloGloGlyGyCysGloGloThrSerGloThrGlyArgAsp 40
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       GlnAlaProGlnGlySerArgSerLeuThrProCysThrCysGlySerSerAspLeuTyr
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              41 LysAsnGlnValGluGlyGluValGlnIleValSerThrAlaThrGlnThrPheLeuAla
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          LeuValThrArgHisAlaAspValIleProValArgArgArgGlyAspSerArgGlySer
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Sequence 6, Application US/10133133A
Publication No. US20030114385A1
GENERAL INFORATION:
APPLICANT: CATHERS, Brian
APPLICANT: NEOTEDON, SASKIA
APPLICANT: STEPARD, Michael
TITLE OF INVENTION: VIRAL ENZYME ACTIVATED PROTOTOXOPHORES
                                                                                                                                                                                                                                 1185
128
129
129
   Complexes
                                                                                                                                                                                                                              Length:
Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                          US-09-965-594-16 (1-197) x US-09-965-594-4 (1-588)
Protease: Inhibitor
                                                                                                                                                                                                                                                                                                           Gaps:
           FILE REFERENCE: DB17Sequences
CURRENT APPLICATION NUMBER: US/09/965,594
CURRENT FILING DATE: 2001-09-27
PRIOR FILING DATE: 1090-01-08
NUMBER OF SEQ ID NOS: 26
SEQ ID NO 4
LENGTH: 588
                                                                                                                                                                                                                                4.84e-100
942.00
94.42%
93.91%
                                                                                                                                                     TYPE: DNA ORGANISM: Hepatitis C virus
                                                                                                                                                                                                                                                              Percent Similarity:
Best Local Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          US-10-133-133A-6
                                                                                                                                                                                                                   Alignment Scores:
```

534

120

100

294

234

80

174 9

```
ValThrargHisAlaaspValIleProValArgArgArgGlyAspSerArgGlySerLeu 141
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   142 LeuSerProArgProIleSerTyrLeuLysGlySerSerGlyGlyProLeuLeuCysPro 161
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          82 ProLysGlyProValThrGlnMetTyrThrAsnValAspLysAspLeuValGlyTrpGln 101
                                                                                                                                                                                                                                                                                                                             22 AlaGlnGlnThrArgClyGluGluGlyCysGlnGluThrSerGlnThrGlyArgAspLys 41 [[[[[[[[[[[[[]]]]]]]]]]]]
                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CysileAsnGlyValCysTrpThrValTyrHisGlyAlaGlyThrArgThrIleAlaSer 81
                                                                                                                                                                                                                                                                                                                GlySerValValIleValGlyArgIleAsnLeuSerGlyAsp-----ThrAlaTyr
TITLE OF INVENTION: AND USE OF SAME TO TREAT VIRAL INPECTIONS FILE REFERENCE: NB 2021.00
CURRENT APPLICATION UNMER: US/10/133,133A
CURRENT FILING DATE: 2002-04-26
PRIOR APPLICATION NUMBER: 60/286,983
PRIOR FILING DATE: 2001-04-27
NUMBER OF SEQ ID NOS: 21
SOFTWARE: FASISEQ for Windows Version 4.0
SEQ ID NO 6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Length:
Matches:
Conservative:
Mismatches:
Indels:
Gaps:
                                                                                                                                                                                                                                                                                       US-09-965-594-16 (1-197) x US-10-133-133A-6 (1-612)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             02-09-742-659-3
: Sequence 3, Application US/09742659
: Patent No. US20010034019A1
GENERAL INFORMATION:
APPLICANT: Hong, Zhi
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Butkiewicz, Nancy J.
                                                                                                                                                                                              1.13e-95
904.50
92.82%
91.28%
87.65%
                                                                                                                                             ORGANISM: Hepatitus C. Virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Zhong, Weidong
Ingravallo, Paul
                                                                                                                                                                                                                       Percent Similarity:
Best Local Similarity:
                                                                                                                                                                                   Alignment Scores Pred. No.:
                                                                                                                                                           US-10-133-133A-6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                APPLICANT:
APPLICANT:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            APPLICANT:
                                                                                                                                  TYPE: DNA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                      62
                                                                                                                                                                                                                                                 Query Match:
                                                                                                                                                                                                            Score:
                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        硆
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                 õ
                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                      õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ద
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
```

```
3354 CGTAGGGCCAGGAGATACTGCTTGGACCAGCCGACGGAATGGTCTCCAAGGGGTGGAGG 3413
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     111::: 111:11 | 111:11 | 111:11 | 111:11 | 111:11 | 111:11 | 111:11 | 111:11 | 111:11 | 111:11 | 111:11 | 111:11 | 111:11 | 111:11 | 111:11 | 111:11 | 111:11 | 111:11 | 111:11 | 111:11 | 111:11 | 111:11 | 111:11 | 111:11 | 111:11 | 111:11 | 111:11 | 111:11 | 111:11 | 111:11 | 111:11 | 111:11 | 111:11 | 111:11 | 111:11 | 111:11 | 111:11 | 111:11 | 111:11 | 111:11 | 111:11 | 111:11 | 111:11 | 111:11 | 111:11 | 111:11 | 111:11 | 111:11 | 111:11 | 111:11 | 111:11 | 111:11 | 111:11 | 111:11 | 111:11 | 111:11 | 111:11 | 111:11 | 111:11 | 111:11 | 111:11 | 111:11 | 111:11 | 111:11 | 111:11 | 111:11 | 111:11 | 111:11 | 111:11 | 111:11 | 111:11 | 111:11 | 111:11 | 111:11 | 111:11 | 111:11 | 111:11 | 111:11 | 111:11 | 111:11 | 111:11 | 111:11 | 111:11 | 111:11 | 111:11 | 111:11 | 111:11 | 111:11 | 111:11 | 111:11 | 111:11 | 111:11 | 111:11 | 111:11 | 111:11 | 111:11 | 111:11 | 111:11 | 111:11 | 111:11 | 111:11 | 111:11 | 111:11 | 111:11 | 111:11 | 111:11 | 111:11 | 111:11 | 111:11 | 111:11 | 111:11 | 111:11 | 111:11 | 111:11 | 111:11 | 111:11 | 111:11 | 111:11 | 111:11 | 111:11 | 111:11 | 111:11 | 111:11 | 111:11 | 111:11 | 111:11 | 111:11 | 111:11 | 111:11 | 111:11 | 111:11 | 111:11 | 111:11 | 111:11 | 111:11 | 111:11 | 111:11 | 111:11 | 111:11 | 111:11 | 111:11 | 111:11 | 111:11 | 111:11 | 111:11 | 111:11 | 111:11 | 111:11 | 11:11 | 111:11 | 111:11 | 111:11 | 111:11 | 111:11 | 111:11 | 111:11 | 111:11 | 111:11 | 111:11 | 111:11 | 111:11 | 111:11 | 111:11 | 111:11 | 111:11 | 111:11 | 111:11 | 111:11 | 111:11 | 111:11 | 111:11 | 111:11 | 111:11 | 111:11 | 111:11 | 111:11 | 111:11 | 111:11 | 111:11 | 111:11 | 111:11 | 111:11 | 111:11 | 111:11 | 111:11 | 111:11 | 111:11 | 111:11 | 111:11 | 111:11 | 111:11 | 11:11 | 111:11 | 111:11 | 111:11 | 111:11 | 111:11 | 111:11 | 11:11 | 111:11 | 111:11 | 111:11 | 111:11 | 111:11 | 111:11 | 111:11 | 111:11 | 111:11 | 111:11 | 111:11 | 111:11 | 111:11 | 111:11 | 111:11 | 111:11 | 111:11 | 111:11 | 11:11 | 11:11 | 11:11 | 11:11 | 11:11 | 11:11 | 11:11 | 11:11 | 11:11 | 11:11 | 11:11 | 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             3474 ACCAGTCTGACTGGCCGGGACAAAACCAAGTGGAGGGTGAGGTCCAGATCGTGTCAACT 3533
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            134 ArgGlyAspSerArgGlySerLeuLeuSerProArgProlleSerTyrLeuLysGlySer 153
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               154 SerGlyGlyProLeuLeuCysProAlaGlyHisAlaValGlyIlePheArgAlaAlaVal 173
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   15 ---LeuSerGlyAspThrAlaTyrAlaGlnGlnThrArgGlyGluGluGlyCysGlnGlu
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             34 ThrSerGlnThrGlyArgAspLysAsnGlnValGluGlyGluValGlnIleValSerThr
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              54 AlaThrGlnThrPheLeuAlaThrCysIleAsnGlyValCysTrpThrValTyrHisGly
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             9646
172
10
13
9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Length:
Matches:
Conservative:
Mismatches:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   US-09-965-594-16 (1-197) x US-09-742-659-3 (1-9646)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              3 LysLysGlySerValValIleValGlyArglleAsn
                                                                                                 ON: Chimeric HCV/GBV-B viruses
ID01116
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Indels:
                                                                                                                                                                   CURRENT APPLICATION NUMBER: US/09/742,659
CURRENT FILING DATE: 2000-12-21
PRIOR APPLICATION NUMBER: US 60/171,469
PRIOR FILING DATE: 1999-12-22
NUMBER OF SEQ ID NOS: 16
SOFTWARE: PATENTIN VET. 2.1
SEQ ID NO 3
APPLICANT: Wright-Minoque, Jacquelyn
APPLICANT: Lau, Johnson Y.
APPLICANT: Lemon, Stanley M.
TITLE OF INVENTION: Chimeric HCV/GBV-
FILE REPERENCE: ID01116
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         2.78e-92
888.50
89.22%
84.31%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        194 MetArgSerPro 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                   ; TYPE: DNA
; ORGANISM: Hepatitis
US-09-742-659-3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Percent Similarity:
Best Local Similarity:
Query Match:
DB:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                    9646
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Pred. No.:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Score:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           임
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
```

RESULT 11 US-09-238-076-1

```
174 CysThrArgGlyValAlaLysAlaValAspPheIleProValGluSerLeuGluThrThr 193
                      AspLysAspLeuValGlyTrpGlnAlaProGlnGlySerArgSerLeuThrProCysThr 113
                                                                                                  CysGlySerSerAspLeuTyrLeuValThrArgHisAlaAspValIleProValArgArg 133
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             TITLE OF INVENTION: FUNCTIONAL DNA CLONE FOR HEPATITIS C VIRUS (HCV) AND USES THEREOF
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC Compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/95,937
FILING DATE: 28-No. US20030028010A1-2001
CLASSIFICATION: <u >CURROWN></u>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  NUMBER OF SEQUENCES: 21
CORRESPONDENCE ADDRESS:
ADDRESSE: HOWELL & HAFERKAMP, L.C.
STREET: 7733 FORSYTH BLVD., SUITE 1400
CITY: ST. LOUIS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             PRIOR APPLICATION DATE:

APPLICATION NUMBER: US/09/034,756
FILING DATE: 04-May-1998
ATTORNEY/AGENT INFORMATION:
NAME: HOLLAND, DONALD R.
REGISTRATION NUMBER: 35,197
REFERENCE/POCKET NUMBER: 6029-4831
TELECOMMUNICATION INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SEQUENCE DESCRIPTION: SEQ ID NO: 1:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             et al
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 TELEPHONE: 314-727-5188
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               LENGIH: 9646 base pairs
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Sequence 1, Application US/0995937; Publication No. US20030028010A1
GENERAL INFORMATION: APPLICANT: RICE, CHARLES et
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               TELEFAX: 314-727-6092 INFORMATION FOR SEQ ID NO: 1:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              TYPE: nucleic acid
STRANDEDNESS: double
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          COMPUTER READABLE FORM:
                                                                                                                                                                                                                                                                                                                                                                                                                                       3954 ATGAGATCCCCG 3965
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       TOPOLOGY: linear MOLECULE TYPE: CDNA
                                                                                                                                                                                                                                                                                                                                                                                                                  194 MetArgSerPro 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     COUNTRY: USA
ZIP: 63105
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              HYPOTHETICAL: NO
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     STATE: MO
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ANTI-SENSE: NO
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 US-09-995-937-1
                      94
                                                                                                    114
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              RESULT 12
                                                                                                                                                                                                                                                                                                                                                                                                                                                       셤
                                                   8
                                                                                                  à
                                                                                                                                   q
                                                                                                                                                                            ò
                                                                                                                                                                                                                 셤
                                                                                                                                                                                                                                                          ô
                                                                                                                                                                                                                                                                                             셤
                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                           ద
                                                                                                                                                                                                                                                                                                                                                                                                                  à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           3594 GCCGGAACGAGGACCATCGCATCACCCAAGGGTCCTGTCATCCAGATGTATACCAATGTG 3653
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          3354 CGTAGGGGCCAGGAGATACTGCTTGGGCCAGCCGACGGAATGGTCTCCAAGGGGTGGAGG 3413
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    3414 TIGCIGCCCCATCACGGCGTACGCCCAGCAGACGAGAGGCCTCCTAGGGTGTATAAIC 3473
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     AlaGlyThrArgThrIleAlaSerProLysGlyProValThrGlnMetTyrThrAsnVal 93
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         54 AlaThrGlnThrPheLeuAlaThrCysIleAsnGlyValCysTrpThrValTyrHisGly 73
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               15 ---LeuSerGlyAspThrAlaTyrAlaGlnGlnThrArgGlyGluGluGlyCysGlnGlu 33
                                                                   TITLE OF INVENTION: FUNCTIONAL DNA CLONE FOR HEPATITIS OF TITLE OF INVENTION: FUNCTIONAL DNA CLONE FOR HEPATITIS OF TITLE OF INVENTION: VIRUS (HCV) AND USES THEREOF NUMBER OF SEQUENCES: 21
CORRESPONDENCE ADDRESS:
ADDRESSEE: HOWELL & HAPERKAMP, L.C.
STREET: 7733 FORSYTH BLVD., SUITE 1400
CITY: ST. LOUIS
STATE: MO
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      9646
172
10
13
9
                                                                                                                                                                                                                                                      ZIP: 63105
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IRM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/238,076
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Length:
Matches:
Conservative:
Mismatches:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              US-09-965-594-16 (1-197) x US-09-238-076-1 (1-9646)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   3 LysLysGlySerValVallleValGlyArglleAsn-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ATTORNEY AGENT INFORMATION:
NAME: HOLLAND, DONALD R.
REGISTRATION NUMBER: 35,197
REFERENCE/DOCKET NUMBER: 6029-4831
TELECOMMUNICATION INFORMATION:
TELEPHONE: 314-727-5188
TELEFAX: 314-727-6092
                                                                                                                                                                                                                                                                                                                                                                                                                                                 PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 09/034,756
FILING DATE:
                                                             APPLICANT: RICE, CHARLES et al.
Sequence 1, Application US/09238076
Patent No. US20020102540A1
GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         LENGTH: 9646 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  TELEFAX: 314-727-6092
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         888.50
89.22%
84.31%
86.09%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       MOLECULE TYPE: CDNA HYPOTHETICAL: NO
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Percent Similarity:
Best Local Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                       CLASSIFICATION:
                                                                                                                                                                                                                                                                                                                                                                                                                  FILING DATE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ANTI-SENSE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   US-09-238-076-1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     74
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Query Match:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Pred. No.:
```

ô g à 셤 ö 셤 δ a õ

```
Percent Similarity:
Best Local Similarity:
                                                                                                                                                                                                                                                                                                Alignment Scores:
                                                                                                                                                                                                                                                                           US-09-917-563-1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               3534
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  3654
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              3714
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          3774
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      3834
                                                                                                                                                                                                                                                                                                                                                                                                                                                    12
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             94
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        114
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    134
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                154
                                                                                                                                                                                                                                                                                                                                                    Query Match:
DB:
                                                                                                                                                                                                                                                                                                           Pred. No.:
                                                                                                                                                                                                                                                                                                                       Score:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     셤
                                                                                                                                                                                                                                                                                                                                                                                                         ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                           a
                                                                                                                                                                                                                                                                                                                                                                                                                                                   à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              a
                                                                                                                                                                                                                                                                                                                                                                                                                              3833
                                                                                                                                                                                                                                                                                                                                                                                                                                                   3773
                                                                                                                                                                                                                                             3654 GACCAAGACCTIGIGGGCIGGCCCGCICCTCAAGGIICCCGCTCAIIGACACCCIGCACC 3713
                                                                                                                       3354 CGTAGGGGCCAGGAGATACTGCTTGGGCCAGCCGACGGAATGGTCTCCAAGGGGTGGAGG 3413
                                                                                                                                                       3474 ACCAGCCTGACTGGCCGGGACAAAACCAAGTGGAGGGTGAGGTCCAGATCGTGTCAACT 3533
                                                                                                                                                                                                                                                                                                                                                                                                       AspLysAspLeuValGlyTrpGlnalaProGlnGlySerArgSerLeuThrProCysThr 113
                                                                                                                                                                                                                                                                                                                                                                CysGlySerSerAspLeuTyrLeuValThrArgHisAlaAspValIleProValArgArg 133
                                                                                                                                                                                                                                54 AlaThrGlnThrPheLeuAlaThrCysIleAsnGlyValCysTrpThrValTyrHisGly 73
                                                                                                                                                                                                                                                                           74 AlaGlyThrArgThrIleAlaSerProLysGlyProValThrGlnMetTyrThrAsnVal 93
                                                                                                                                            ---LeuSerGlyAspThrAlaTyrAlaGlnGlnThrArgGlyGluGluGlyCysGlnGlu 33
                                                                                                                                                                                                                                                                                                                                                                           ThrSerGlnThrGlyArgAspLysAsnGlnValGluGlyGluValGlnIleValSerThr
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   TITLE OF INVENTION: FUNCTIONAL DNA CLONE FOR HEPATITIS C VIRUS (HCV) AND USES THEREOF
9646
172
10
13
9
 Length:
Matches:
Conservative:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CORRESPONDENCE ADDRESS:
ADDRESSEE: HOWELL & HAFERKAMP, L.C.
STREET: 7733 FORSYTH BLVD., SUITE 1400
                                                                                                  LysLysGlySerValVallleValGlyArglleAsn--
                                                                            US-09-965-594-16 (1-197) x US-09-995-937-1 (1-9646)
                                 Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          et al
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Sequence 1, Application US/09917563
Publication No. US20030073080A1
GENERAL INFORMATION:
APPLICANT: RICE, CHARLES et
2.78e-92
888.50
89.22%
84.31%
86.09%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                COMPUTER READABLE FORM:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        NUMBER OF SEQUENCES: 21
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              3954 ATGAGATCCCCG 3965
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CITY: ST. LOUIS
STATE: MO
COUNTRY: USA
ZIP: 63105
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         MetArgSerPro 197
                    Percent Similarity:
Best Local Similarity:
Query Match:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                US-09-917-563-1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                         3834
                                                                                                  ٣
                                                                                                                                             15
                                                                                                                                                                                       34
                                                                                                                                                                                                                                                                                                                     96
                                                                                                                                                                                                                                                                                                                                                                114
                                                                                                                                                                                                                                                                                                                                                                                                                                                    154
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         194
  Pred. No.:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     RESULT 13
            Score:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         à
                                                                                                   õ
                                                                                                                       g
                                                                                                                                                                 a
                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                             8
                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                         셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 õ
                                                                                                                                             ò
                                                                                                                                                                                       ò
                                                                                                                                                                                                          g
                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                           õ
                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                     ò
```

```
SerGlyGlyProLeuLeuCysProAlaGlyHisAlaValGlyIlePheArgAlaAlaVal 173
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              54 AlaThrGlnThrPheLeuAlaThrCysIleAsnGlyValCysTrpThrValTyrHisGly 73
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             3414 TIGCIGCCCCATCACGCCTACGCCCAGCAGACGAGGCCTCCTAGGGIGTATAATC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CysGlySerSerAspLeuTyrLeuValThrArgHisAlaAspValIleProValArgArg
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ArgGlyAspSerArgGlySerLeuLeuSerProArgProIleSerTyrLeuLysGlySer
PatentIn Release #1.0, Version #1.30
                                                                                                                                                                                                                                                                                                                                                                                                                                9646
172
10
13
9
                                                                                                                                                                                                                                                                                                                                                                                                                                  Length:
Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          3 LysLysGlySerValVallleValGlyArglleAsn-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              US-09-965-594-16 (1-197) x US-09-917-563-1 (1-9646)
                                                                                                                                                                       6029-4831
           CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/917,563
FILING DATE: 27-Jul-2001
CLASSIFICATION: <Unknown>
                                                                                 FALLY AFFLLIALION JUNERS. 09/238,076
FILING DATE: 26-JAN-1999
ATTORNEY AGENT INFORMATION:
NAME: HOLLAND, DONALD R.
REGISTRATION NUMBER: 35,197
REFERENCE/DOCKET NUMBER: 6029-483
TELECOMMUNICATION INFORMATION:
TELEFON: 314-72-6092
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CRARACTERISTICS:
LENGTH: 9646 base pairs
                                                                                                                                                                                                                                                                                                          TOPOLOGY: linear
MOLECULE TYPE: CDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
SEQUENCE DESCRIPTION: SEQ ID NO: 1:
                                                                                                                                                                                                                                                                            TYPE: nucleic acid
STRANDEDNESS: double
                                                                                                                                                                                                                                                                                                                                                                                                                                  2.78e-92
888.50
89.228
84.318
                                                                           PRIOR APPLICATION DATA:
```

3473

3533

23

3713

3773

133

à g ò g 133

113

83

```
3834 TCGGGGGGTCCGCTGTTGTGCCCCGGGGACACGCCGTGGGCCTATTCAGGGCCGCGGGG 3893
                                                                                                                                                                                                                                                                                                                                                                                                                   SerGlyGlyProLeuLeuCysProAlaGlyHisAlaValGlyIlePheArgAlaAlaVal 173
                                                   74 AlaGlyThrArgThrIleAlaSerProLysGlyProValThrGlnMetTyrThrAsnVal
                                                                                                                                                                                                                                                                        114 CysGlySerSerAspLeuTyrLeuValThrArgHisAlaAspValIleProValArgArg
                                                                                                                                                                                                  94 AsplysAspleuValGlyTrpGlnAlaProGlnGlySerArgSerLeuThrProCysThr
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ublication mo. ....
GENERAL INFORMATION:
APPLICANT: RICE, CHARLES et al.
TITLE OF INVENTION: FUNCTIONAL DNA CLONE FOR HEPAIITIS C
VIRUS (HCV) AND USES THEREOF
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     COMPUTER: IBM PC compatible OPERATING SYSTEM: PC-DOS/MS-DOS SOFTWARE: Patentin Release $1.0, Version $1.30
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/995,937
FILING DATE: 28-No. US20030028010A1-2001
CLASSIFICATION: <Unknown>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CORRESPONDENCE ADDRESS:
ADDRESSEE: HOWELL & HAFERKAMP, L.C.
STREET: 7733 FORSYTH BLVD., SUITE 1400
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     NAME: HOLLAND, DONALD R.
REGISTRATION NUMBER: 35,197
REGISTRATION NUMBER: 6029-4831
TELECOMMUNICATION INFORMATION:
TELEPRAS: 314-727-5188
TELEFAX: 314-727-5092
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                APPLICATION NUMBER: US/09/034,756 FILING DATE: 04-May-1998
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      LENGIH: 12980 base pairs
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   MEDIUM TYPE: Floppy disk
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Sequence 5, Application US/09995937 Publication No. US20030028010A1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ATTORNEY/AGENT INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SEQUENCE CHARACTERISTICS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 INFORMATION FOR SEQ ID NO: 5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              COMPUTER READABLE FORM:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                PRIOR APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  3954 ATGAGATCCCCG 3965
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CITY: ST. LOUIS
STATE: MO
COUNTRY: USA
21P: 63105
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                MetArgSerPro 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        US-09-995-937-5
                                                                                                                                                                                                                                                                                                                                                                                                                     154
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                194
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        RESULT 15
                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                               õ
                                                                                     g
                                                                                                                           ö
                                                                                                                                                        셤
                                                                                                                                                                                                ö
                                                                                                                                                                                                                                    a
                                                                                                                                                                                                                                                                        õ
                                                                                                                                                                                                                                                                                                                                               õ
                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  3414 ITGCTGGCGCCCATCACGGCGTACGCCCAGCAGACGAGGGCGTCCTAGGGTGTAAATC 3473
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             3354 CGTAGGGGCCAGGAGATACTGCTTGGGCCAGCCGACGGATGGTCTCCAAGGGGTGGAGG 3413
                   CysthrArgGlyValAlaLysAlaValAspPheIleProValGluSerLeuGluThrThr 193
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ---LeuSerGlyAspThrAlaTyrAlaGlnGlnThrArgGlyGluGluGlyCysGlnGlu 33
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ThrSerGlnThrGlyArgAspLysAsnGlnValGluGlyGluValGlnIleValSerThr
                                                                                                                                                                     Sequence 5, Application US/09238076
Patent No. US20020102540A1
GENERAL INFORMATION:
APPLICANT:
TITLE OF INVENTION: FUNCTIONAL DNA CLONE FOR HEPATITIS C
TITLE OF INVENTION: VIRUS (HCV) AND USES THEREOF
NUMBER OF SEQUENCES: 21
CORRESPONDENCE ADDRESS:
ADDRESSEE: HOWELL & HAFERKAMP, L.C.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SYSTEM: PC-DOS/MS-DOS
PatentIn Release #1.0, Version #1.30
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      US-09-965-594-16 (1-197) x US-09-238-076-5 (1-12980)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           LysLysGlySerValVallleValGlyArgIleAsn
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Length:
                                                                                                                                                                                                                                                                                                                        SEE: HOWELL & HAFERKAMP, L.C.
1: 7733 FORSYTH BLVD., SUITE 1400
ST. LOUIS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Gaps:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 6029-4831
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 09/034,756
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/238,076
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ATTORREY/AGENT INFORMATION:
NAME: HOLLAND, DONALD R.
RECISTRATION NUMBER: 35,197
REFERENCE/OCKET NUMBER: 6029
TELECOMUNICATION INFORMATION:
TELEPHONE: 314-727-5188
TELEFANCE 314-727-6092
                                                                                                                                                                                                                                                                                                                                                                                                                                  COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    TELEFAX: 314-727-6092 INFORMATION FOR SEQ ID NO: 5:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SEQUENCE CHARACTERISTICS:
LENGTH: 12980 base pairs
TYPE: nucleic acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         4.06e-92
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          888.50
89.228
84.318
86.098
                                                                                                            3954 ATGAGATCCCCG 3965
                                                                        194 MetArgSerPro 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          STRANDEDNESS: double
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        OPERATING SYSTEM:
SOFTWARE: Patenti
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                MOLECULE TYPE: CDNA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 linear
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                      USA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Percent Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 HYPOTHETICAL:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        FILING DATE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    FILING DATE
                                                                                                                                                                                                                                                                                                                                                                                                                       ZIP: 63105
                                                                                                                                                                                                                                                                                                                                                                                     ç
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 TOPOLOGY:
                                                                                                                                                                                                                                                                                                                                                                                   STATE: MC COUNTRY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   US-09-238-076-5
                                                                                                                                                                                                                                                                                                                                               STREET:
CITY: S1
                                                                                                                                                                 US-09-238-076-5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ٣
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               15
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      34
 174
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Query Match:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Pred. No.:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Score:
```

g ô

a ŝ

```
3354 CGTAGGGGCCAGGAGATACTGCTTGGGCCAGCCGACGGAATGGTCTCCAAGGGGTGGAGG 3413
                                                                                                                                                                                                                               3414 TIGCIGGCGCCCAICACGGCGIACGCCCAGCAGACGAGAGGCCICCTAGGGIGIAIAAIC 3473
                                                                                                                                                                                                                                                                                                                                                                                                                                  134 ArgGlyAspSerArgGlySerLeuLeuSerProArgProIleSerTyrLeuLysGlySer 153
                                                                                                                                                                                                                                                                                                                       74 AlaGlyThrArgThrIleAlaSerProLysGlyProValThrGlnMetTyrThrAsnVal 93
                                                                                                                                                                         15 ---LeuSerGlyAspThrAlaTyrAlaGlnGlnThrArgGlyGluGluGlyCysGlnGlu 33
                                                                                        Length:
Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                                                        US-09-965-594-16 (1-197) x US-09-995-937-5 (1-12980)
                                                       Ś
                                                        ..
Q
                                                        SEQUENCE DESCRIPTION: SEQ ID
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: CDNA
HYPOTHERICAL: NO
ANTI-SENSE: NO
                                                                                         4.06e-92
888.50
89.22%
84.31%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       11111111111
3954 ATGAGATCCCCG 3965
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            194 MetArgSerPro 197
                                                                                                      Percent Similarity:
Best Local Similarity:
Query Match:
DB:
                                                                                   Alignment Scores:
                                                                US-09-995-937-5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                         ò
                                                                                                                                                                                           a
                                                                                                                                                                                                                              a
                                                                                                                                                                                                                                                                q
                                                                                                                                                                                                                                                                                     à
                                                                                                                                                                                                                                                                                                     a
                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                        qq
                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                           đ
                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ã
                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                            õ
                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
```

Search completed: August 31, 2003, 04:54:21 Job time : 190.482 secs

```
GenCore version 5.1.6
```

Copyright (c) 1993 · 2003 Compugen Ltd.	ıs_p2n model
1993 - 2003	ing frame_plu
Copyright (c)	ic search, us
	protein - nucleic search, using frame_plus_p2n model
	_

August 30, 2003, 19:20:43; Search time 1910.31 Seconds

Run on: 8

(without alignments)
2506.388 Million cell updates/sec 1 MKKKGSVVIVGRINLSGDTA......VAKAVDFIPVESLETIMRSP 197 22781392 seqs, 12152238056 residues BLOSUM62 Xgapop 10.0 , Xgapext Ygapop 10.0 , Ygapext Fgapop 6.0 , Fgapext Delop 6.0 , Delext US-09-965-594-16 1032 Title: Perfect score: Scoring table: Sequence: Searched:

Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0 Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL-frame+ p2n.model -DEV-x1p
-MODEL-frame+ p2n.model -DEV-x1p
-MODEL-frame+ p2n.model -DEV-x1p
-Gord\_1/USPTO\_spool/US09955594/runat\_29082003\_151919\_28322/app\_query.fasta\_1.2872
-DB-EST -OFMT-fastap -SUFFIX-rst -MINMATCH-0.1 -LOOPCL-0 -LOOPEXT-0
-UNITS-b1ts -START-1 -END--1 -MATRIX-blosum62 -TRANS-human40.cdi -LIST-45
-DOCALIGN-200 -THR\_SCORE-PCT -THR\_MAX-100 -THR\_MIN-0 -ALIGN-15 -MODE-LOCAL
-UNFNT-PCO -NORM-ext -HEA.PSIZE-500 -MILEN-0 -MAXLEN-200000000
-USER-USO955594\_eCGN\_1 1\_12630\_erunat\_29082003\_151919\_28322 -NCPU-6 -ICPU-3
-NO\_MARP -LARGEQUERY -NEC\_SCORES-0 -WAIT -DSPLOCK-100 -LONGLOG
-DEV\_INDOUT-120 -WARN\_TIMEOUT-30 -THRADS-1 -XGAPOP+10 -XGAPEXT-0.5 -FGAPOP-6
-FGAPEXT-7 -YGAPOP-10 -YGAPEXT-0.5 -DELEXT-7

em\_estba:\* EST:\* Database :

em\_esthum:
em\_esthum:
em\_esthum:
em\_esthu:
em\_estpl:
em\_htc:
gb\_estl:
gb\_htc:
gb\_estl:
gb\_htc:
gb\_estl:
gb\_estl:
gb\_estl:
gb\_estl:
gb\_estl:

em\_estfun:\* em\_estom:\* em\_gss\_hum:\* em\_gss\_inv:\* em\_gss\_pln:\* 

em\_gss\_pro:\* em\_gss\_rod:\* em\_gss\_mus:\* em\_gss\_phg:\* em\_gss\_vrl:\* gb\_gssl:\*

em\_gss\_vrt:\*

em\_gss\_fun:\* em\_gss\_mam:\*

29: gb\_gss2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

	scription	926101 AGENCOUR	531437	392487	BF182274 601804028	304699	42842	380545 Mus mus	01625 B	24121 BJ02412	16176	20999	3398	93316	5803	244	33	280	713	416	886	343	637	75	995	BE289911 601089126	1821	34	615	9608	486	1657	132	1584	1992	1748	964	200	45	620249	525	46602	33286 HQ01M02	08007 UI-E-	5435 AL856435	25317 K-ESTO	
CHINGRADO		BQ92610	BF631437	248	BF182274	691	82	24	162	112	517	99	339	331	8	324	3039	728	3771	3341	888	2034	663	5424	660	BE289911	882	834	3761	904	9848	3465	3513	5858	S.	2374	7406	975	974	AQ620249	3462	4660	3328	<b>470800</b>	85643	253	
	gth DB	403 1	_	_	846 10			<b>⊣</b>	-				_	٦.	-		_	<b>α</b>	_		٠,	٦.	- ·	ص .			_	6	-	_	-	_	_	_	٦.		~	_	-	_	7	_	7	-	σ	_	
ø	Query Match Length	9.6	9.	ت	ı.	'n.	₹.	₹.	۳.	m.	e.	m.	7.	7.	7.	٦.		Γ.	٦.	7.	0.	0	0.	0.	٥.	٥. ه	on.	σ.	6.	σ.	σ.	6.	σ.	<b>æ</b> .	æ.	<b>0</b> 0.	æ.	€.	œ.	8.	ω.	∞.	۲.	8.7	۲.	۲.	
	Score	01.	99.5	æ	86	٠,	97.5	97	96	96	96	96	95	95	95	<b>7</b> 6	94	9	93.5	m'	66	σ,	92.5	ni (	٠, '	250	20		•	•	91.5	•	•	91	91	91	91	91	σ	90.5	·	o.	90	90	90	06	
	Result No.		7	m	υ •				<b>ထ</b> ပ				12		7	15	c 16	17	18	19	C) I	c 21	710	~ (	C 74	N (	. 1	c 27	7	c 29	m	31	c 35	33	34	32	36	37	c 38	39				-	C 44	4	

J403 bp mRNA linear EST 20-AUG-2002 AGENCOURT\_8752655 NIH\_MGC\_130 Mus musculus cDNA clone IMAGE:6335718 B0926101 Mus musculus (house mouse)
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalla; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Hus.
1 (bases 1 to 1403) BQ926101.1 GI:22341132 RESULT 1 BQ926101/c LOCUS DEFINITION VERSION KEYWORDS SOURCE ORGANISM ACCESSION

REFERENCE

```
library
                                                                                                                                                                                                         EST
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SOUTCE
                                                                                                                     DEFINITION
                                                                                                                                                                                                                                         ORGANISM
                                                                                                                                                                                                                                                                                                         REFERENCE
AUTHORS
                                                                                                                                                                       ACCESSION
                                                                                                                                                                                       VERSION
KEYWORDS
SOURCE
                                                                                                                                                                                                                                                                                                                                                                                                                         JOURNAL
                                                                     RESULT 2
BF631437
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               FEATURES
                                                                                                                                                                                                                                                                                                                                                                         TITLE
                                                                                                                                                                                                                                                                                                                                                                                                                                         COMMENT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1378 GGGTGTGTTCANCGGTCAGGACAGGTCGCC---GCACACTCGACGCGTCGGCCAGAGACT 1322
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1321 TGTCGGGGGCGCGTTGGCGCATACCCCGGGTCGGATCGAGGTCAGCGCCGCTTGTATACA 1262
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1249 GGGGTA---TGGTTATCACGGGCTGGGCAGGTACT------TCCCCTAAAGCG 1205
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1204 GCGCCGTGCCGAGTATATATCCGGCGAGTGCGAAGCGCACGGGGCGTGGAACGTTGACC 1145
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                261 GAGGGAAA------CAG 1250
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1087 GGCACGGTGATGGTTACTACGCGGCAGACCGGCTCCAACGGCGCTCTCTAACAGACGC 1028
                                                                                                                                                                                                                                                                84 lyProvalThrGlnMetTyrThrAsnValAspLysAspLeuValGlyTrpGlnAlaProG 104
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         104 InGlySerArgSerLeuThrProCysThrCysGlySerSerAspLeuTyrLeuValThrA 124
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        11 GlyArgileAsnLeuSerGlyAspThrAlaTyrAlaGlnGlnThrArgGlyGluGluGly 30
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            65 GlyValCysTrpThrValTyrHisGlyAlaGlyThrArgThrIleAlaSerProLys--G 84
                                                                 Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Mark Maconochie, Ph.D. and Nancy L. Freeman,
Ph.D.
                                                                                                              CDNA Library Preparation: ResGen, Invitrogen Corp

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Agencourt Bioscience Corporation

Clone distribution: MGC clone distribution information can be

found through the.I.M.A.G.E. Consortium/LLNL at:

http://lmage.llnl.gov

Plate: LLMAN1398 row: i column: 07

High quality sequence stop: 101.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          31 CysGlnGluThr-----SerGlnThrGlyArgAspLysAsnGln-----Val
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            45 GluGlyGluValGlnIleValSerThrAlaThrGlnThrPheLeuAlaThrCysIleAsn
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     137 ----SerArgGlySer-----LeuLeuSerProArgProlleSerTyrLeuLysG
                     National Institutes of Health, Mammalian Gene Collection (MGC)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                124 rgHisAlaAspValIleProValArgArgArgGlyAsp----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         44
62
44
43
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Length:
Matches:
                                                   Contact: Robert Strausberg, Ph.D.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          US-09-965-594-16 (1-197) x BQ926101 (1-1403)
   NIH-MGC http://mgc.nci.nih.gov/.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      15.5
101.50
36.75%
29.52%
9.84%
                                     Unpublished
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Percent Similarity:
Best Local Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                      297
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Query Match:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Pred. No.:
AUTHORS
TITLE
JOURNAL
COMMENT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                       BASE COUNT
                                                                                                                                                                                                                                                     FEATURES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ORIGIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ô
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
```

```
112 1/5ersec1/901 156

11 11111

997 GARAGCCCCCCT 954

12 GARAGCCCCCCCT 954

12 GARAGCCCCCCCT 954

13 GARAGCCCCCCCT 954

14 GARAGCCCCCCCT 954

15 GARAGCCCCCCCT 954

16 GARAGCCCCCCCCT 954

16 GARACTCC 10 113092107

17 GARAGCCCCCCCCT 1110psida; Debales; Doacese; Pooldes

16 GARACTC 10 113092107

17 GARACTC 10 1110psida; Debales; Doacese; Pooldes

16 GARACTC 10 113092107

17 GARACTC 10 1110psida; Dales; Doacese; Pooldes

17 GARACTC 10 1110psida; Dales; Doacese; Pooldes

18 GARACTC 10 113092107

18 GARACTC 10 1110psida; Dales; Doacese; Pooldes

18 GARACTC 10 1110psida; Dales; Dales
```

FEATURES Location/Qualifiers Source 1. 1199   //organism="#nono sapiens" //mol_type="mRNA"   //mol_type="m	/clone="IMAGE:6192708" //sex=male" /tissue.type="sympathetic trunk" /dev_stage="adult, 16 yr" /lab.host="abillo" /lab.host="lungs"; sympathetic trunk"	/note="vector: ppgw.SpoRT6 (Life Technologies); Site_1: Not1; Site_2: Sal1; cDNA made by oligo-dr priming. Directionally cloned using the following adaptors: 5'-TCGACCCACGCGTCG-3' and 5'-GACTACATCATCGGGAGCGGCCCCT(15)-3'. Size selected >	l kb for average insert length 1.9 kb. This is a primary library, non-amplified. Library constructed by Life Technologies and donated by J. Lupski, M.D./Ph.D. (Baylor College of Medicine); available through Life	BASE COUNT 255 a 362 c 343 g 211 t 28 others ORIGIN	Alignment Scores: 24.8 Length: 1199 Pred: No.: 98.50 Matches: 40 Score: 98.50 Matches: 40 Percent Similarity: 36.77% Conservative: 17	Mismatches: Indels: Gaps:	16 (1-197) x BQ892487 (1-1199) TrpThrValTyrHisGlyAlaGlyThrArgThrIleAlaSerProLysGlyProValThr 87	*0* ISSUED TO THE SECOND TO THE SECOND TO THE SECOND TO THE SECOND	DD 544 CITCCATTACCACATGTGACAGTTT	Oy 119 LeuTyr-LeuValThrArgHisAlaAspValIleProValArg 132 :::	Qy 133	Qy 143SerProArgProlleSerTyrLeuLysGlyserSerGl 155	Oy 155 yGlyProLeuLeuCysProAlaGlyHisAlaValGlyIlePheArgAlaAlaValCysTh 175 1	Oy 175 rArgGlyValAlaLysAlaValAspPheIleProValGluSer 189 :::	RESULT 4 BF182274/c LOCUS BF182274  BF182274  BF182274  DEFINITION 601804028F1 NCI_CGAP_Mam5 Mus musculus cDNA clone IMAGE:4035102 5', mRNA sequence. ACCESSION BF182274
(http://wheat.pw.usda.gov/ggpages OUNT 147 a 216 c 311 g 105 t ent Scores: 11.3 Length:	Score: 99.50 Matches: 47  Percent Similarity: 39.874 Conservative: 14  Best Local Similarity: 30.724 Mismatches: 63  Query Match: 10 Gaps: 6  DB:	US-09-965-594-16 (1-197) x BF631437 (1-779)  Qy	Oy 69 ThrValTyrHisGlyAlaGlyThrArgThrIleAlaSerProLysGlyProValThr 87 	Qy 88GlnMetTyrThrAsnValAspLysAspLeuValGlyTrpGln 101 ::    bb 109 CCCTGTTCTGCCAGGTGGCCGGGTGGACGCGCGCAGGGGAGGAGGAGGCTCGTGGTCG 168	Qy 102 AlaProGlnGlySerArgSerLeuThrProCysThrCysGlySerSer 117 :::	Oy 118 -AspLeuTyrLeuValThrargHisAlaAspValIleProValArgar 133	Oy 133 9ArgGlyAspSerArgGlySerLeuLeuSerProArgProIleSerTyrLeuLySGlySe 153	Qy 153 rSerGly-GlyProLeuLeuCysProAla	Oy 165 laValGlyllePheArgalaValCysThrArg 176	RESULT 3 ROS92487 LOCUS LOCUS DEFINITION AGENCOUFF.8417538 Lupski, sympathetic_trunk Homo sapiens CDNA Clone TWAGE, 5197708 5, mbhs section	ACCESSION BQ892487 VERSION BQ892487.1 GI:22284501 KEYWORDS EST. SOURCE Homo sapiens (human)	_	ACTHORS NIH-MGC http://mgc.ncl.nih.gov/. TITLE National Institutes of Health, Mammalian Gene Collection (MGC) JOUNAL Dapublished COMMENT Contact: Robert Stransbarg Db D		DNA Sequencing by: Agencourt Bioscience Corporation Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov Plate: LLAM13595 row: c column: 13 High quality sequence start: 57

Oy 176 gGlyValAlaLys 180  RESULT 5  RESULT 5  RESULT 5  RESULT 5  RESULT 5  RESULT 6 01888252F1 NIH_MGC_17 Homo:  ACCESSION BF304699.1 GI:11251586  VERSION BF304699.1 GI:11251586  KEYWORDS ST.  SOURNEL Homo saplens (human)  ORGANISM Homo saplens (human)  ORGANISM Homo saplens (human)  ORGANISM Homo saplens (human)  AUTHORS HOTHP://mgc.nci.nlh.go.  TITLE NALHONAL INSLITUTES Of Healtl  JOURNAL CONTACT: RObert Strausberg, IEmail: Gapabs-remail: ATCC  COMMENT CONTACT: RObert Strausberg, IEmail: Gapabs-remail: ATCC  CONMENT CONTACT: ROBER Strausberg, IEmail: Gapabs-remail: ATCC  CONMENT CONTACT: ROBER STRAUSDERGY, IEMAIL: Gapabs-remail: ATCC  CONMENT CONTACT: ROBER STRAUSDERGY, IEMAIL: Gapabs-remail: ATCC  CONMENT CONTACT: ROBER STRAUSDERGY, IEMAIL: GAPACT  CONTACT: CONTACT: ROBER STRAUSDERGY, IEMAIL: GAPACT  CONTACT: ROBER STRAUSDERGY, IEMAIL: GAPACT  CONTACT: CONTACT  CONTACT: ROBER STRAUSDERGY, IEMAIL: GAPACT  CONTACT: ROBER STRAUSDERGY, IEMAIL: GAPACT  CONTACT  CONTACT	/tissue_rippe="rhabd" /tissue_rippe="rhabd" /lab_nost="Dist" = Thabd
BF182274.1 GI:11060416 EST.  Mus musculus (house mouse)  Mus musculus (house mouse)  Mus musculus  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  Mammalia; Eutheria; Rodentia; Sciurognathi: Muridae; Murinae; Mus.  1 (Pases I to 846)  Onpublished  Contact: Robert Strausberg, Ph.D.  Email: cgapbs-r@mail.nih.gov/.  Tissue Procurement: Lothar Hennighausen Ph.D., Robin Humphreys  Contact: Robert Strausberg, Ph.D.  Email: cgapbs-r@mail.nih.gov/.  Tissue Procurement: Lothar Hennighausen Ph.D., Robin Humphreys  CDNA Library Preparation: Life Technologies, Inc.  CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)  Clone distribution: MGC clone distribution information can be  found through the I.M.A.G.E. Consortium/LLNL at:  http://laage.llnl.gov  Plate: LLAM9308 row: g column: 07  High quality sequence stop: 696.  Location/Qualiflers  Location/Qualif	: : ty: arity: arity: yalaG 11:: 11:: 11:: 11:: 11:: 12:: 12:: 13:: 13
VERSION KEYWORDS SOURCE ORGANISH ORGANISH TITLE JOURNAL COMMENT SOUFCE SOUFCE	BASE COUNT ORIGIN Alignment Scores Pred. No.: Score: Score: Score: Score: Guery Match: DB: US-09-965-594-16 UP-09-965-594-16 UP-09-965-594-16 UP-09-09-965-594-16 UP-09-09-965-594-16 UP-09-09-09-09-09-09-09-09-09-09-09-09-09-

```
habdomyosarcoma*
habdomyosarcoma*
B (phage-resistant)*
USCL17
USCL217

14 bp mRNA linear EST 21-NOV-2000 sapiens cDNA clone IMAGE:4122276 5',
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      serLeuThr---ProCysThrCysGlySerSerAsp 118
::
tCGAGAAGAGGACCGTGTACCTGC------ 599
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               GATGGTGTCCAAGCTCGCGCTGCTACTGGAAGT 494
                                                                                                                                                                                                                                                                                                                                                                                     a; Craniata; Vertebrata; Euteleostomi; s; Catarrhini; Hominidae; Homo.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Ling Hong/Rubin Laboratory
The I.M.A.G.E. Consortium (LLNL)
Genomics, Inc.
Lone distribution information can be
E. Consortium/LLNL at: image.llnl.gov
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               .gov/.
ith, Mammalian Gene Collection (MGC)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        984
32
55
22
52
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Length:
Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Ph.D.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       piens"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Gaps:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          984)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    .90
```

61		H & H & H &	VURDAYANGE AKOB0545  AKOB0545  AKOB0545  FINITION Mus musculus 7 days neonate cerebellum cDNA, RIKEN full-length enriched 1lbrary, clone:A730082L10 product:weakly similar to zinc finger protein (fragment) [Mus musculus], full insert sequence.  CESSION AKOB0545. I GI:26348600  WWCDDS HTC: CAP trapper.  ONCEDS Mus musculus (house mouse)  ONCANISM Mus musculus (house mouse)  ONCANISM Mus musculus (house mouse)  ONCANISM Mus musculus (house mouse)  AMD Mus musculus (house mouse)  AMD Mus musculus (house mouse)  AMD Mus musculus (house mouse)  ONCANISM Mus musculus (house mouse)  AND Mus musculus (house mouse)  AUTHORS Latentyota; Metazoa: Chordata; Craniata; Vertebrata; Euteleostomi;  Musmmalia; Eutheria; Rodentia; Sciurognathi; Muridae; Musinae; Mus.  AUTHORS Carninci, P. and Hayashizaki, Y.  TITLE High-efficiency full-length cDNA cloning  MUDLINE 99279253  PUBNED 10349636	
* da * da	6 6 6 6 6 6	Oy Oy Oy Oy Oy Ox Ox	ACCESSION VERYODS SOURCE ORGANISM REFERENCE ORGANISM REFERENCE AUTHORS JOURNAL MEDLINE PUBLED PUBLED	REFERENCE AUTHORS TITLE JOURNAL MEDLINE PUBMED REFERENCE AUTHORS TITLE
Oy 154 SerGlyGlyProLeuLeuCysProAlaGlyHisAlaValGlyIlePheArgAlaAlaVal 173	o z	Underty Contact: Robert Strausberg, Ph.D.  Contact: Robert Strausberg, Ph.D.  Email: cgapbs-rémail.nih.gov		Alignment Scores:  Alignment Scores:  Alignment Scores:  1.5

JOURNAL

```
/protein_id="BAC37940.1"
/db_xref="GI:26348601"
/db_xref="GI:26348601"
/translation="DSCIPERASPGRSILTPRGDGFFIKEKLSAARAVGPGPSVAFGVS
TRYGAQAQFGQRRVGACRREGECLSRKPRRQHVPPVGPHYYGLSGGRRIPPPAGE
AQAAGBARQQYPHPGRPGRYVPPQGAAGLIPALAARQVPVGPVRGREGPRRAPRHS
PRPVPTALGFSFGQGGPAPPLLAPANGRSVGLAL"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 BJ001625
BJ001625 MF01SSA CDNA Oryzias latipes CDNA clone MF01SSA025C02 5',
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Oryzias latipes (Japanese medaka)
Oryzias latipes
Cytzias latipes
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      120 rLeuValThrArgHisAlaAspValIleProValArg---ArgArgGlyAspSerArgGl 139
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      100 pGlnAlaProGlnGlySerArgSerLeuThrProCysThrCysGlySerSerAspLeuTy 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    411 -crescascresceaastrecretresacesceatrassessessessesaceaasses 469
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      172 aValCysThrArgGlyValAlaLysAlaVal-----AspPheIleProValGluSerLe 190
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              646
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    81 Ser-ProLysGlyProValThrGlnMetTyrThrAsnValAspLysAspLeuValGlyTr 100
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         139 ySerLeuLeuSerProArgProlle-------SerTyrLeuLysGlySe 153
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         153 rSerGlyGlyProLeuLeuCysProAla -- - GlyHisAlaValGlyIlePheArgAlaAl 172
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 /note="unnamed protein product; putative
weakly similar to zinc finger protein (fragment)
musculus] (PIR[148722, evidence: FASTY, 50.7%ID,
57.6%length, match-601)"
/codon_start="
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    338 AGIGCCGCATCCGCCTGGGCGGCCACATGGAACAGTGGTG----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       35
8
11
11
35
35
8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Length:
Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                            u
                                                                                                                                                                                                                                                                                                                                                                                                                       228
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Gaps:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               US-09-965-594-16 (1-197) x AK080545 (1-1141)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 TrpThrValTyrHisGlyAlaGly----
                                                                                                                                                                                                                                                                                                                                                                                                                            ø
                                                                                                                                                                                                                                                                                                                                                                                                                       353
                                                                                                                                                                                                                                                                                                                                            /note="putative"
                                                                                                                                                                                                                                                                                                                                                                                               /note-"putative"
316 c 353
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   BJ001625.1 GI:17364516
                                                                                                                                                                                                                                                                                                                       1118. .1123
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 190 uGluThrThrMet 194
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 647 GCCAGGTACACTT 659
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       32.5
97.00
39.318
31.728
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      mRNA sequence.
BJ001625
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Percent Similarity:
Best Local Similarity:
Query Match:
                                                                                                                                                                                                                                                                                                                     polyA_signal
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                EST
                                                                                                                                                                                                                                                                                                                                                                       polyA_site
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 99
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    378
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   RESULT 8
BJ001625/c
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   LOCUS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ..
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ORGANISM
                                                                                                                                                                                                                                                                                                                                                                                                                       BASE COUNT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ACCESSION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      VERSION
KEYWORDS
SOURCE
                  CDS
                                                                                                                                                                                                                                                                                                                                                                                                                                                  ORIGIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           d
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ρp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    QY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    원
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    õ
                                                                                                       Arawaw, T., Shinaqawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y., Arawawa, T., Hara, A., Fukunishi, Y., Konno, H., Adachi, J., Fukuda, S., Arawawa, T., Hara, A., Fukunishi, Y., Kiyosawa, H., Kondo, S., Yamanaka, I., Saito, T., Caito, T., Casaki, Y., Gojobori, T., Bono, H., Kasukawa, T., Saito, R., Kadota, K., Matsuda, Y., Gissi, C., King, B., Kochiwa, H., Fleischmann, W., Gasterland, T., Gissi, C., King, B., Kochiwa, H., Kuehl, P., Lewis, S., Matsuo, Y., Nikaido, I., Pesole, G., Tomita, M., Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H., Baldarelli, R., Barsh, G., Blakk, J., Boffelli, D., Bojunga, N., Baldarelli, R., Bronstein, M.J., Bult, C., Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D., Hofmann, M., Hume, D. A., Kamiya, M., Lee, N. H., Lyons, P., Nordone, P., Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H., Toyo-oka, K., Schonbach, C., Seya, T., Shibata, Y., Storch, K., Schonbach, C., Seya, T., Shibata, Y., Storch, K., Schonbach, C., Seya, T., Shibata, Y., Kawaji, H., Kohtsuki, S., and Hayashizaki, Y., Ramashizaki, Y., Rawashizaki, Y., Rawashiza
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               /tissue_type="cerebellum"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="7 days neonate"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. DIvision of Experimental Animal Research in Riken contributed to
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Group Phase I & II Team.
Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         The FANTOM Consortium and the RIKEN Genome Exploration Research
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Functional annotation of a full-length mouse cDNA collection
Nature 409 (6821), 685-690 (2001)
21085660
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   prepare mouse tissues.
Please visit our web site for further details.
URL:http://genome.gsc.riken.go.jp/
URL:http://fantom.gsc.riken.go.jp/.
1. 1141
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              /db_xref-"FANTOM_DB:A730082L10"
/db_xref-"taxon:10090"
Genome Res. 10 (11), 1757-1771 (2000)
20530913
11076861
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           /organism-"Mus musculus"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        /clone-"A730082L10"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                /mol_type="mRNA"
/strain="C57BL/6J"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Fax:81-45-503-9216)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         11217851
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       JOURNAL
REFERENCE
AUTHORS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               TITLE
JOURNAL
MEDLINE
PUBMED
                                                                                          REFERENCE
AUTHORS
                                       MEDLINE
PUBMED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 TITLE
JOURNAL
```

REFERENCE AUTHORS

TITLE

FEATURES

COMMENT

```
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neoplerygii; Teleostei; Euteleostei; Neoteleostei; Acanthomorpha; Acanthopterygii; Percomorpha; Atherinomorpha; Beloniformes; Acanthopterygii; Percomorpha; Atherinomorpha; (bases I to 643)
Kohara,Y., Shin-1,T., Kimura,T., Narita,T., Jindo,T. and Takeda,H., Medaka Est Project in Takeda's lab
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   96 AspLeuValGlyTrpGlnAlaProGlnGlySerArgSerLeuThrProCysThrCysGly 115
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Arg-----ProlleSerTyrLeuLysGlySerSerGlyGlyProLeuLeuCysPro 161
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            242 AAAAATGACGTAGAACCAAAAGACACAGATCCAACCACACATGTTCTGGTTCTACGGGCT 301
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   -------TGTTGGAGAACCTATCACAGTTCCTGCTTTAGAGCAACGGCA 343
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              344 GCTCTGCGGCGGCGAGGAGCTCCTGGGCCAGTTGTGACTCGT--------385
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         -----GGAGGACGAAGAGCGTCACCCCGGAGCTGTAGGCTGCAGGGATGCGGATGTGGC 439
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            SerSerAspLeuTyrLeuValThrArg-----124
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 440 TCTGCT-----TTGGTTCCTCGTCCTCGGATCATCTTCCTCACCTGACCTTCCA 490
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       125 HisAlaAspValileProValArgArgArgGlyAspSerArgGlySerLeuLeuSerPro 144
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          491 CATCCAGGIGICGCCAGCGCIGGTCIGACGGGIGAIGGGAGAGGCCGGACAGGCAGCAGT 550
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     551 CGGGGGTGAATCTCTGCAGGACGTCTTCACGGCGGATCAGGAGGACCGACTCGCTGCAGA 610
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                61 ThrCyslleAsnGlyValCysTrpThrValTyrHisGlyAlaGlyThrArgThrIleAla 80
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         81 SerProLys------GlyProValThrGlnMetTyrThrAsnValAspLys 95
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         41 LysAsnGlnValGluGlyGluValGlnIleValSerThrAlaThrGlnThrPheLeuAla
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             /clone="MF01SSA143D12"
/sex="mixture of female and male"
/tissue_type="whole embryo"
/dev_stage="segmentation stage 20 - 25"
/clone_lib="MF01SSA cDNA"
148 c 148 g 176 t
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    643
7 4 3
9 6
9 6
6
                                                                                                                                                                                                                                                                                   Center For Genetic Resource Information
National Institute of Genetics
1111 Yata, Mibhima, Shizuoka 411-8540, Japan
Tat: 81-559-81-6856
Fax: 81-559-81-6855
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Conservative:
Mismatches:
                                                                                                                                                                                                                                                                                                                                                                                                               1. .643
/organism="Oryzias latipes"
/mol_type="mRNA"
/strain="Hd-rR"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Length:
Matches:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Indels:
                                                                          Oryzias latipes (Japanese medaka)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    US-09-965-594-16 (1-197) x BJ024121 (1-643)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               /db_xref-"taxon:8090"
                                                                                                                                                                                                                                                                                                                                                                               Email: tshini@genes.nig.ac.jp.
Location/Qualifiers
                     BJ024121
BJ024121.1 GI:17377389
                                                                                                                                                                                                                                               Unpublished
Contact: Tadasu Shin-i
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    19.4
96.00
34.72%
29.86%
9.30%
                                                                                                Oryzias latipes
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Percent Similarity:
Best Local Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           171
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Alignment Scores:
                                                          EST
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   302
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         386
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              116
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               145
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Query Match:
DB:
                                                                                                                                                                                                                                                                                                                                                                                                                        source
                                                                                                ORGANISM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Pred. No.:
                   ACCESSION
VERSION
KEYWORDS
SOURCE
                                                                                                                                                                                        REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           BASE COUNT
                                                                                                                                                                                                                                                                                                                                                                                                   FEATURES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ORIGIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   BJ024121 BF01SSA CDNA OFYZIAS latipes CDNA clone MF01SSA143D12 3',
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          -------TGTTGGAGAACCTATCACAGTTCCTGCTTTAGAGCAACGGCA 410
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              96 AspLeuValGlyTrpGlnAlaProGlnGlySerArgSerLeuThrProCysThrCysGly 115
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  111:::
313 TCTGCT------TTGGTTCCTCGTCCTCGGATCATCTTCCTCACCTGACCTTCCA 263
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             HisAlaAspValIleProValArgArgArgGlyAspSerArgGlySerLeuLeuSerPro 144
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 -----GGAGGACGAAGAGCGTCACCCCGGAGCTGTAGGCTGCAGGGATGCGGATGTGGC 314
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SerSerAspLeuTyrLeuValThrArg----- 124
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     262 CATCCAGGTGTCGCCAGCGCTGGTCTGACGGGTGATGGGAGAGGCCGGACAGGCAGTGT 203
                                       1 (bases 1 to 615)
Kohara, Y., Shin-i, T., Kimura, T., Narita, T., Jindo, T. and Takeda, H.
Medaka EST Project in Takeda's lab
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    LysAsnGlnValGluGlyGluValGlnIleValSerThrAlaThrGlnThrPheLeuAla 60
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            61 ThrCysIleAsnGlyValCysTrpThrValTyrHisGlyAlaGlyThrArgThrIleAla 80
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       81 SerProLys------GlyProvalThrGlnMetTyrThrAsnValAspLys 95
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Acanthomorpha; Acanthopterygii; Percomorpha; Atherinomorpha;
Beloniformes; Adrianichthyidae; Oryziinae; Oryzias.
                                                                                                                                                                                                                                                                                                                                                                           /sex="mixture of female and male"
/tissue_type="whole embryo"
/dev_stage="segmentation stage 20 - 25"
/clone_lib="MFOLSSA cDNA"
166 c 165 g 144 t
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 615
43
7
58
36
6
                                                                                                            Contact: Tadasu Shin-i
Contact: Tadasu Shin-i
Center For Genetic Resource Information
National Institute of Genetics
Ini Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856
Fax: 81-559-81-6856
Exa: 81-559-81-6856
Email: tshini@enes.nig.ac.jp.
Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Matches:
Conservative:
Mismatches:
                                                                                                                                                                                                                                                                                 /organism="Oryzias latipes"
/mol_type="mRNA"
/strain="Hd-rR"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Length:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Gaps:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  US-09-965-594-16 (1-197) x BJ001625 (1-615)
                                                                                                                                                                                                                                                                                                                                        /db_xref="taxon:8090"
/clone="MF01SSA025C02"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               18.3
96.00
34.728
29.86%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     162 AlaGlyHisAla 165
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        GCCTCTGCTGCA 131
                                                                                                                                                                                                                                                             1, .615
                                                                                              Unpublished
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Percent Similarity:
Best Local Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         116
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               125
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  451
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 367
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             BJ024121
LOCUS
DEFINITION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Query Match:
                                                                                                                                                                                                                                                                   source
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Alignment S
Pred. No.:
                                     REFERENCE
AUTHORS
TITLE
JOURNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                        BASE COUNT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               RESULT 9
                                                                                                                                                                                                                                               FEATURES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ORIGIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ద
```

us-09-965-594-16.rst

Oy 125 HisAlaAspValIleProValArgArgGlyAspSerArgGlySerLeuLeuSerPro 144	RESULT 11 CB95099 LOCUS DEFINITION AGENCOURT_13445496 NIH_MGC_177 Mus musculus cDNA clone NACESSION CB950999 1 GI:30205777 RESION CB950999 1 GI:30205777 REYWORDS SOURCE Mus musculus (house mouse) ORGANISM Mus musculus (house mouse) REFERENCE 1 (bases 1 to 1031) REFERENCE 1 (bases 1 to 1031)	NIH-W NIH-W NIH-W Unpub Conta Conta CDNA CDNA CLODA Found http: High	/mol_type="mrNa" /db_xref='taxon:10090" /db_xref='taxon:10090" /db_host="lhiGE:3016.62" /lab_host="bh10B (T1-phage-resistant)" /clone_lib="nik_MGC_177" /clone_lib="nik_MGC	Alignment Scores:  Alignment Scores:  Pred. No.:  Score:  Score:  Percent Similarity:  Percen
16 61 0N		Ill Yata, Mishina, Shizuoka 411-6540, Japan Tel: 81-559-81-6856 Fax: 81-559-81-6855 Fax: 81-559-81-6855 Email: tshini@genes.nig.ac.jp. Location/Qualifiers  1. 754 //organism="O	Alignment Scores:  Pred. No.: Pred. No.: Score: Score: Score: Score: Standarity: 34.728 Conservative: Percent Similarity: 9.308 Conservative: 12 Conservative: 13 Conservative: 14 Conservative: 15 Conservative: 16 Conservative: 17 Conservative: 18 Conservative: 18 Conservative: 18 Conservative: 19 Conservative: 19 Conservative: 10 Conservative:	Db   302

US-09-965-594-16 (1-197) x CA728398 (1-580)   GO AlaThrCysIleAsnGlyValCysTrpThrValTyrHisGlyAlaGlyThrArg 77	116 SerSerAsp-LeuTyrLeuValThrArgHisAlaAspValIleProVal	<u> </u>	AUTHORS NIH-MGC http://mgc.nci.nih.gov/.  TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  JORNAL Uppublished  Contact: Robert Strausberg, Ph.D.  Email: cgapbs-r@mail.nih.gov  Tissue Procurement: ATCC  CONM. Library Preparation: Ling Hong/Rubin Laboratory  CDNA Library Parayaration: Ling Hong/Rubin Laboratory  CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  DNA Sequencing by: Incyte Genomics, Inc.  Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov  Plate: LLCM965 row: 1 column: 03	High quality sequence stop: 637.  FEATURES Location/Qualifiers  Source /organism="Homo sapiens" /mol_type="mkRA" /db_xref="taxon:9606" /clone="traxon:9606" /lissue_type="thabdomyosarcoma" /lab_host="bh108 (phage resistant)" /clone_lib="NHIM (phage resistant)" /clone_lib="NHIM (phage resistant)" /clone_lib="NHIM (phage resistant)" /sice_2: xhoi; cDNA maded by oligo-dT priming. Site_2: xhoi; cDNA maded by oligo-dT priming. Directionally cloned into EcoRl/xhoi sites using the following 5' adaptor: GGCACGAG(G). Size-selected by for average insert size 1:8tb. Library constructed by for average insert size 1:8tb. Library constructed by for average insert size 1:8tb. Library constructed by	OUNT 23 ent Scores: No.: t Similarity
60 - SD	6 8 8	RESULT : BF203310 LOCUS DEFINITY ACCESSIC VERSION KEYWORD SOURCE ORGAN	AUTHO TITLE JOURN COMMENT	FEAT	BASE CONIGIN Alignm Pred. Score: Percen
64 AsnGlyValCysTrpThrValTyrHisGlyAlaGlyThrArgThrIleAlaScrProLys 83	604ACCACCCATGGGGCTTGTTGTTCCCCGCCTCCCGGGGCAATTACA 651  138 gGlySerLeuLeuSerProArgprolleSerTyrLeuLyGlySerSerGly 155  11	Db 712 GTCCCCCTTTTGTTACCCACACTTTGTGGGA 748  RESULT 12 CA728398 LOCUS CA728398 CA728308 CA728308 CA728308 CA728308 CA728308 CA728308 CA7283	EMERYORA, VITIGIPANITAE: SITEPIDOPNYIA; EMDIYOPNYIA; ENGREDOPNYIA; EMARYORA; Tracheopnyia; Spermatophyta; Magnollophyta; Liliopsida; Poales; Poaceae; Pooldeae; Triticeae; Troman,	Fax: 302-631-2607  Email: Scott.V.ThngeyeUSA.dupont.com Seq primer: M13. Location/Qualifiers  1. 580	Inment Scores: 21.2 Length: 5 is 95.00 Matches: 3 is 95.00 Matches: 3 is 1.00 Matches: 3 is 1.00 Matches: 3 is 1.00 Matches: 3 is 1.00 Marches: 3 is 1.00 Matches: 3

larity: 34, imilarity: 24, 9.2	594-16 (1-197) x BM915803 (1-1146) 22 AlaglaGlaThrArgGlyGluGluGlyCysGlnGluThrSer-GlnThrGlyArgAsply 41		41 sAsnGlnValGluGlyGluValGlnIleValSerThrAlaThrGlnThrPheLeuAlaTh 61 ::: ::::::: :::::::::::::::::::::::	CINCACACACACACACACACACACACACACACACACACAC	Arghrieses::::::::::::::::::::::::::::::::::	948 AGCGGAGGGGCGCCGCGTCGGGGCGCGGGGGGGGGGGGCCCG	107 gSerLeuThrProCysThrCysGlySerSerAspLeuTyrLeuValThrArgHisAlaAs 127 	<pre>pValileProValArgArgArgGlyAspSerArgGlySerLeuLeuSerProArgProil       ::</pre>	833	yllepheArgAlaAlaValCysThrArgGlyValAlaLysAlaValAspPheIlePr	741 GGCCTTCCGGCGTCTGGTGTCTTCGCGGTCTGCCGGGGGG		-		CONTROL: CALIES MADEL DUNE BOX 91000 Duke University Durham, NC 27708-1000 Tel: 919 613 8159 Fax: 919 613 8177 Email: chauser@duke.edu.
Percent Simi Best Local S Query Match: DB:	US-09-965-594-16		ò d			a o	λο QΩ	ζŎ	ag ov d	à à	අ රු අ	RESULT 15 BF863244 LOCUS DEFINITION	ACCESSION VERSION KEYWORDS SOURCE ORGANISM	TITLE	
Best Local Similarity: 40.58%       Mismatches: 27         Query Match: 9.21%       Indels: 4         DB: 10       Gaps: 1         US-09-965-594-16 (1-197) x BF203316 (1-961)		TGTGGAGCTCAGAGATGCTGACTCTGAGTGCAGGAGGGGACAGAGATGCGGATCGGA	133 gArgGlyAspSerArgGlySerLeuLeuSer-ProArgProIleSerTyrLeuLysGlyS 153 :	153 erSerGlyGlyProLeuLeuCySProAlaGlyHisAlaValGlyIlePheArgAlaAlaV 173 	173 alcysThrArgGlyvalAlaLys 180 	RESULT 14 HM915803/c LOCOS BM915803 1146 bp mRNA linear EST 12-MAR-2002 DEFINITION AGENCOURT_6639455 NIH_MGC_41 Homo sapiens CDNA clone IMAGE:5482056	5', mRNA sequence. ACCESSION BM915803 VERSION BM915803.1 GI:19366182 KRYWORDS FST.	W.	Mammalia, Eutherial Primates; Catarrhini; Hominidae; Homo. Nammalia, Eutherial 1 to 1146) AUTHORS NIH-MGC http://mgc.nci.nih.gov/. TITLE National institutes of Health, Mammalian Gene Collection (MGC)		CDNA Library Preparation: Rubin Laboratory CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: Agencourt Bioscience Corporation Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:		<pre>source 11146 //organism="Homo sapiens" //mol_type="mRNA" /db_xref="taxon:9606" /clone="IMAGE:5482056" //tissue_type="amelanotic melanoma, cell line" /lab host="DH10B (phage-resistant)" /clone_lib="NIH_MGC_41" /note="Organ: skin; Vector: porB7; Site_2: /note="Organ: skin; Vector: porB7; Site_2:</pre>	into EcoRI/Abol sites using the following 5 adaptor: GGCACGAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-ODNA synthesis kit (Strategene) and Supersoript II RT (Life Technologies). Note: this is a	BASE COUNT       169 a       492 c       344 g       141 t         ORIGIN       Alignment Scores:       51.3       Length:       1146         Pred: No.:       95.00       Matches:       45

```
1...'Organism="Chiamydomonas reinhardtii"

/mol_type="mRNA"
/mol_type="mRNA"
/strain="Cc-1690 wild type mt+ 21gr"
/db_xref="taxon:3055"
/dbone_lib="C. reinhardtii CC-1690, Stress condition I,
normalized, Lambda Zap II"
/note="Vector: pBluescript II SK-; Site_l: ECORI; Site_2:
/note="Vector: pBluescript II SK-; Site_l: Ahr); TAP-S (30 min,
lhr, 4hr); TAP-P (4hr, 12hr, 24hr), NO3 to NH4 (30min, lhr,
/note="Vector: pBluescript II SK-; Side, Intro lambda and cDNA synthesized.
/note="Vector: pBluescript II SK-; Dlasmids were excised from the lambda
/pBluescript II SK- plasmids were excised from the lambda
/AP clones by superinfection with ExAssist (Stratagene)
/phage: The library was normalized using method 4 described
in Bonaldo et al (1996) Genome Research 6: 791-806."
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  87 ThrGlnMetfyrThrAsnValAspLysAspLeuValGlyTrpGlnAlaProGlnGlySer 106
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   231 ACAAAGTTACATACACGG-----AAGGACCAGCGCGTTGGCCACCCCCTTGGAGCCG 284
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           107 ArgSerLeuThrProCysThrCysGlySerSerAspLeuTyrLeuValThrArgHisAla 126
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           127 Aspvalile ----- 138
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 :::::||||
345 GAGATCATITGCATGIGGCTTTAGICACCCCAAGAAGAGCCTGGGAGIGGGCATITATAA 404
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         -----GlySerLeuLeuSerProArgProIleSerTyrLeu 150
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 405 GAAGGGGACGGGAATTCCGTTTGCGGAAAAGTGAGCGCCCCAAGGTCTGACCAAGTGCTA 464
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       71 TyrHisGlyAlaGlyThrArgThr11eAlaSerProLys-----GlyProVal 86
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            701
32
14
45
45
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          151 LysGlySerSerGlyGlyProLeuLeuCysProAlaGly 163
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  165 CTCCAAGGCAGCAATGGGAGCCTTTCGCGGTGTGCGGGT 503
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Length:
Matches:
Conservative:
Mismatches:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Gaps:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 US-09-965-594-16 (1-197) x BF863244 (1-701)
    Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            33.9
94.00
40.71%
28.32%
9.11%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Percent Similarity:
Best Local Similarity:
Query Match:
DB:
                                                                                                                                                                                                                                                                                                                                                                                                                                                           173
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Aliqument Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         139
                                                                                                                                                                                                                                                                                                                                                                                                                                                         BASE COUNT
ORIGIN
                           source
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Pred. No.:
FEATURES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Score:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             QQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  쉽
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       q
```

Search completed: August 31, 2003, 04:27:34 Job time: 1916.31 secs

```
Hepatitis C virus
                                                                                                              August 30, 2003, 17:42:58 ; Search time 44.6227 Seconds (without alignments) 700.745 Million cell updates/sec
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           /SIDSI/gcdata/geneseq/geneseqp-embl/AA1980.DAT:*
/SIDSI/gcdata/geneseq/geneseqp-embl/AA1981.DAT:*
/SIDSI/gcdata/geneseq/geneseqp-embl/AA1981.DAT:*
/SIDSI/gcdata/geneseq/geneseqp-embl/AA1981.DAT:*
/SIDSI/gcdata/geneseq/geneseqp-embl/AA1981.DAT:*
/SIDSI/gcdata/geneseq/geneseqp-embl/AA1985.DAT:*
/SIDSI/gcdata/geneseq/geneseqp-embl/AA1985.DAT:*
/SIDSI/gcdata/geneseq/geneseqp-embl/AA1987.DAT:*
/SIDSI/gcdata/geneseq/geneseqp-embl/AA1987.DAT:*
/SIDSI/gcdata/geneseq/geneseqp-embl/AA1989.DAT:*
/SIDSI/gcdata/geneseqy-geneseqp-embl/AA1989.DAT:*
/SIDSI/gcdata/geneseqy-embl/AA1989.DAT:*
/SIDSI/gcdata/geneseqp-embl/AA1989.DAT:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    | SIDSI/gcddta/geneseq/geneseqreenal/AA1991.DAT:
| SIDSI/gcddta/geneseq/geneseqreenal/AA1992.DAT:
| SIDSI/gcddta/geneseq/geneseqp-enal/AA1992.DAT:
| SIDSI/gcddta/geneseq/geneseqp-enal/AA1993.DAT:
| SIDSI/gcddta/geneseq/geneseqp-enal/AA1995.DAT:
| SIDSI/gcddta/geneseq-geneseqp-enal/AA1995.DAT:
| SIDSI/gcddta/geneseq-geneseqp-enal/AA1996.DAT:
| SIDSI/gcddta/geneseq/geneseqp-enal/AA1999.DAT:
| SIDSI/gcgdta/geneseq/geneseqp-enal/AA1999.DAT:
| SIDSI/gcgdta/geneseq/geneseqp-enal/AA1999.DAT:
| SIDSI/gcgdta/geneseq/geneseqp-enal/AA1999.DAT:
| SIDSI/gcgdta/geneseq/geneseqp-enal/AA2000.DAT:
| SIDSI/gcgdta/geneseq/geneseqp-enal/AA2000.DAT:
| SIDSI/gcgdta/geneseq/geneseqp-enal/AA2000.DAT:
| SIDSI/gcgdta/geneseq/geneseqp-enal/AA2000.DAT:
| SIDSI/gcgdta/geneseq/geneseqp-enal/AA2000.DAT:
| SIDSI/gcgdta/geneseq/geneseqp-enal/AA2003.DAT:
| SIDSI/gcgdata/geneseq/geneseqp-enal/AA2003.DAT:
| SIDSI/gcgdata/geneseq/geneseqp-enal/AA2003.DAT:
| SIDSI/gcgdata/geneseq/geneseqp-enal/AA2003.DAT:
                                                                                                                                                                                                                                       1 MKKKGSVVIVGRINLSGDTA.....VAKAVDFIPVESLETIMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Description
                                                                                                                                                                                                                                                                                                                                                                                   1107863
GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.
                                                                                                                                                                                                                                                                                                                                                                               Total number of hits satisfying chosen parameters:
                                                                                                                                                                                                                                                                                                                                          1107863 seqs, 158726573 residues
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SUMMARIES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      AAB15223
AAB15224
AAB15222
AAB15222
AAB15221
AAB15220
AAB15220
AAB15220
AAB15220
                                                                              - protein search, using sw model
                                                                                                                                                                                                                                                                               BLOSUM62
Gapop 10.0 , Gapext 0.5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 A_Geneseq_19Jun03:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ü
                                                                                                                                                                                                                                                                                                                                                                                                                       length: 0
length: 2000000000
                                                                                                                                                                                        US-09-965-594-18
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          DB
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        197
197
197
197
195
195
665
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Query
Match 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      100.0
98.8
98.5
97.8
93.5
93.0
89.7
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Title:
Perfect score:
Sequence:
                                                                                                                                                                                                                                                                                                                                                                                                                         sed
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1017
1005
1005
995
985
951
946
                                                                                                                                                                                                                                                                               Scoring table:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Score
                                                                                                                                                                                                                                                                                                                                                                                                                       Minimum DB s
Maximum DB s
                                                                              OM protein
                                                                                                                                                                                                                                                                                                                                          Searched:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Database
                                                                                                                      Run on:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Result
```

978.5 86.4 668 877.5 86.3 666 874.5 86.3 666 874.5 86.0 666 874.5 86.0 666 874.5 86.0 666 874.5 86.0 666 873.5 86.0 666 870.5 85.6 67 870.5 85.6 67 870.5 85.6 67 870.5 85.6 67 870.5 85.6 67 870.5 85.6 21 870.5 86.0 22 870.5 86	0 AAY24947 0 AAY24941 0 AAY24942 0 AAY24945 0 AAY24946 0 AAY24946 0 AAY24946 0 AAY27884 0 AAY27884 0 AAY27894 0 AAY27894 0 AAY24944	20         AAX17890           20         AAX17882         HCV NS4A-NS3 compl           20         AAX17883         HCV NS4A-NS3 compl           20         AAX17885         HCV NS4A-NS3 compl           20         AAX17887         HCV NS4A-NS3 compl           20         AAX17885         HCV NS4A-NS3 compl           20         AAX17885         HCV NS4A-NS3 compl           20         AAX17888         HCV NS4A-NS3 compl           21         AAX17888         HCV NS4A-NS3 compl           22         AAW93482         HCV NS4A-NS3 compl           24         AAW93482         HCV NS4A-NS3 compl           24         ABP71460         HCPATILIS C VITUS           24         ABY0128         HCPATILIS C VITUS           24         ABF40120         HCPATILIS C VITUS           24         ABF4012         HCPATICS C VITUS           26         HAX1022         HCPATICS C VITUS           27         AAX9223         HCV Genomic anino           28         AAP92041         SEQUENCE C encoded se           20         AAP92041         SEQUENCE C encoded by           20         AAP92050         SEQUENCE C encoded by           20         AAP92080	ALIGNMENTS  ein; 197 AA.  ry)  NS3 fusion protease #5.  ; viral replication; chronic liver disease;  ncer; mutant; mutein.	00345. 15271. SQUIBB CO. Her S, Zhang Y, Goldfarb V;
	778 778 777 777 777 737 737 737	200 222 22 22 22 22 22 22 22 22 22 22 22	5223 standard; prott 5223; EC-2000 (first ent. titis C virus NS4A-1 titis; NS3 protease r failure: liver cas titis C virus.	00040707-A1.  JUL-2000.  JAN-2000; 2000WO-US  JAN-1999; 99US-01  IM ) BRISTOL-MYERS  tekind M, Weinheim  tekind M, Weinheim  sono-465976/40.

N

```
The present sequence is a mutated version of a fusion protein created using the Hepatitis C virus (HCV) NS3 and NS4A protease enzymes. These proteins are both essential for the replication of the virus, acting to cleave its replicative proteins from the polyprotein produced from the HCV genome. Inhibitors of the two proteins should be effective as antiviral treatments of the two proteins should be effective as chronic liver disease such as cirrhosis, liver failure and liver cancer. The present invention concerns a number of NS3 mutants and NS3-NS4A well as enabling structural studies of the protease and proteins which can be used to identify inhibitors of this type, as well as enabling structural studies of the protease and
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        61 ISINGYLATYYHGAGTATIASPKGPYTQMYTAVDKDLYGWQAPQGGSRSLTPCTCGSSDLY 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         61 TSINGVLWTVYHGAGTRTIASPKGPVTQMYTNVDKDLVGWQAPQGSRSLTPCTCGSSDLY 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        LVTRHADVIPVRRRCDSRGSLLSPRPISYLKGSSGGPLLCPACHAVGIFRAAVSTRGVAK 180
                                                                                                                                                                                                                                                                                                                                                                                                                        substitution of a hydrophobic alpha-helix 0 amino acid to a hydrophilic amino acid, useful for screening inhibitors that may treat hepatitis C
                                                                                                                                                                                                                                                                                                                                                                                                        1 MKKKGSVVIVGRINLSGDTAYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVSTATQTFLA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Hepatitis; NS3 protease; viral replication; chronic liver disease;
liver failure; liver cancer; mutant; mutein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Hepatitis C virus NS4A-NS3 fusion protease #6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  (BRIM ) BRISTOL-MYERS SQUIBB CO.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  AAB15224 standard; protein; 197
                                                                 Claim 23; F1g 15; 66pp; English
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      AVDFIPVESLETTMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       AVDFIPVESLETTMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               06-JAN-2000; 2000WO-US00345
                                                                                                                                                                                                                                                                                                                                                        Best Local Similarity 100.
Matches 197; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  WPI; 2000-465976/40.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   N-PSDB; AAA73333
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 MO200040707-A1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Wittekind M,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 08-JAN-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  19-DEC-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                13-JUL-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Synthetic.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     AAB15224;
                                                                                                                                                                                                                                                                                                           Sequence
                                                                                                                                                                                                                                                                                                                                          Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     RESULT 2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     AAB15224
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                        õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     염
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
 à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ö
```

Ş

(first entry)

Weinheimer S, Zhang Y, Goldfarb V;

99US-0115271

The present sequence is a mutated version of a fusion protein created using the Hepatitis C virus (HCV) NS3 and NS4A protease enzymes. These proteins are both essential for the replication of the virus, acting to cleave its replicative proteins from the polyprotein produced from the HCV genome. Inhibitors of the two proteins should be effective as antiviral treatments of HCV infection. This is useful as HCV can lead to chronic liver disease such as cirrhosis, liver failure and liver cancer. The present invention concerns a number of NS3 mutants and NS3-NS4A fusion proteins which can be used to identify inhibitors of this type, as well as enabling structural studies of the protease and protease: Inhibitor complexes. This sequence contains the alpha-helix0-7 ö 121 LVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVSTRGVAK 180 9 Modified hepatitis C virus (HCV) NS3 protease comprising at least 1 substitution of a hydrophobic alpha-helix 0 amino acid to a hydrophilic amino acid, useful for screening inhibitors that may treat hepatitis C 1 MKKKGSVVIVGRINLSGDTAYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVSTATQTFLA 61 TSINGVLWTVYHGAGTRTIASPKGPVTQMYTNVDKDLVGWQAPQGSRSLTPCTCGSSDLY Hepatitis, NS3 protease, viral replication, chronic liver disease; liver failure, liver cancer; mutant; mutein. ö Length 197; Indels Goldfarb V; Score 1005; DB 21; Pred. No. 2.7e-95; 2; Mismatches 1; Hepatitis C virus NS4A-NS3 fusion protease #4 Wittekind M, Weinheimer S, Zhang Y, Ź AAB15222 standard; protein; 197 8 66pp; English 181 AVDFIPVESLETTMRSP 197 98.8%; 98.5%; (BRIM ) BRISTOL-MYERS SQUIBB 990S-0115271 06-JAN-2000; 2000WO-US00345 (first entry) Best Local Similarity 98.5 Matches 194; Conservative WPI; 2000-465976/40. Hepatitis C virus. Synthetic. 197 AA; Claim 23; Fig 16; N-PSDB; AAA73331 #0200040707-A1 08-JAN-1999; 19-DEC-2000 13-JUL-2000 Seguence AAB15222 Query Match AAB15222 RESULT g Q Q g ŏ ò ò

ó

100.0%; Score 1017; DB 21; Length 197; 100.0%; Pred. No. 1.6e-96; 1ve 0; Mismatches 0; Indels 0;

197 AA;

```
06-JAN-2000; 2000WO-US00345
                                                                                                                                                                                                  19-DEC-2000 (first entry)
                                                                                             Best_Local Similarity 98.5
Matches 194; Conservative
                                                                                                                                                                                                                                                                                        WPI; 2000-465976/40
                                                                                                                                                                                                                             Hepatitis C virus.
Synthetic.
                                                                                 197 AA;
                                                                                                                                                                                                                                         WO200040707-A1
                                                                                                                                                                                                                                                                                 Wittekind M,
                                                                                                                                                                                                                                                                38-JAN-1999;
                                                                                                                                                                                                                                                 13-JUL-2000
                                                                                                                                61
                                                                                                                                                       181
                                                                                                                                                               181
                                                                                                                                                                                           AAB15225;
                                                                                                                        61
                                                                                  Sequence
                                                                                         Query Match
Best Local 3
                                                                          variant
                                                                                                                                                                           RESULT
                                                                                                                                                                                       8
                                                                                                                        ò
                                                                                                                                ద
                                                                                                                                        ŝ
                                                                                                                                               g
                                                                                                                                                       ò
                                                                                                                                                               ద
                                                                                                                                                                                   20
```

```
The present sequence is a mutated version of a fusion protein created using the Hepatitis C virus (HCV) NS3 and NS4A protease enzymes. These proteins are both essential for the replication of the virus, acting to cleave its replicative proteins from the polyprotein produced from the HCV genome. Inhibitors of the two proteins should be effective as antiviral treatments of HCV infection. This is useful as HCV can lead to chronic liver disease such as cirrhosis, liver failure and liver cancer. The present invention concerns a number of NS3 mutants and NS3-NS4A fusion proteins which can be used to identify inhibitors of this type, as well as enabling structural studies of the protease and
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       LVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVSTRGVAK 180
                                                       substitution of a hydrophobic alpha-helix 0 amino acid to a hydrophillic amino acid, useful for screening inhibitors that may treat hepatitis C
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        TSINGVLWTVYHGAGTRTIASPKGPVTQMYTNVDKDLVGWQAPQGSRSLTPCTCGSSDLY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1 MKKKGSVVIVGRINLSGDTAYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVSTATQTFLA
                                      Modified hepatitls C virus (HCV) NS3 protease comprising at least 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Hepatitis; NS3 protease; viral replication; chronic liver disease;
liver failure; liver cancer; mutant; mutein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Score 995; DB 21; Length 197;
Pred. No. 2.9e-94;
2; Mismatches 2; Indels (
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Goldfarb V;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Hepatitis C virus NS4A-NS3 fusion protease #3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Zhang Y,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Ş
                                                                                                                                         Claim 23; Fig 17; 66pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             AAB15221 standard; protein; 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 AVDFIPVESLETTMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         AVDEIPVESLETTMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      97.8%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  (BRIM ) BRISTOL-MYERS SQUIBB
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Weinheimer S,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          99US-0115271
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 06-JAN-2000; 2000WO-US00345
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          19-DEC-2000 (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Best Local Similarity 98.0
Watches 193; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                               197 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Hepatitis C virus.
Synthetic.
N-PSDB; AAA73334
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 WO200040707-A1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Wittekind M,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          08-JAN-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          13-JUL-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                 Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   AAB15221;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Query Match
                                                                                                                                                                                                                                                                                                                                                                                                      variant
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    RESULT 5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                셤
g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ç
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ó
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ö
                                                                                                                                                      The present sequence is a mutated version of a fusion protein created using the Hepatitis C virus (HCV) NG3 and NS4A protease enzymes. These proteins are both essential for the replication of the virus, acting to cleave its replicative proteins from the polyprotein produced from the HCV genome. Inhibitors of the two proteins should be effective as antiviral treatments of the two proteins should be effective as antiviral treatments of the two proteins should be effective as chronic liver disease such as cirrhosis, liver failure and liver cancer. The present invention concerns a number of NG3 mutants and NG3-NS4A fusion proteins which can be used to identify inhibitors of this type, as well as enabling structural studies of the protease and protease:inhibitor complexes. This sequence contains the alpha-helix0-1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             TSINGVLWTVYHGAGTRTIASPKGPVTQMYTNVDKDLVGWQAPQGSRSLTPCTCGSSDLY 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  LVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVSTRGVAK 180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        9
                                    substitution of a hydrophobic alpha-helix 0 amino acid to a hydrophilic amino acid, useful for screening inhibitors that may treat hepatitis C
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1 MKKKGSVVIVGRINLSGDTAYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVSTATQTFLA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Gaps
               Modified hepatitis C virus (HCV) NS3 protease comprising at least 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Hepatitis: NS3 protease; viral replication; chronic liver disease;
liver failure; liver cancer; mutant; mutein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                              Length 197;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Zhang Y, Goldfarb V;
                                                                                                                                                                                                                                                                                                                                                                                                                                                              Score 1002; DB 21;
Pred. No. 5.6e-95;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Hepatitis C virus NS4A-NS3 fusion protease #7
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       0; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Z
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       AAB15225 standard; protein; 197
                                                                                                                      Claim 23; Fig 14; 66pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 AVDFIPVESLETTMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              AVDFIPVESLETTMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                              98.5%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           (BRIM ) BRISTOL-MYERS SQUIBB
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Weinheimer S,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  99US-0115271
```

ö

Gaps

ö

9 9

2000-465976/40

The present sequence is a mutated version of a fusion protein created using the Hepatitis C virus (HCV) NS3 and NS4A protease enzymes. These proteins are both essential for the replication of the virus, acting to cleave its replicative proteins from the polyprotein produced from the HCV genome. Inhibitors of the two proteins should be effective as antiviral treatments of HCV infection. This is useful as HCV can lead to chronic liver disease such as cirrhosis, liver failure and liver cancer. The present invention concerns a number of NS3 mutants and NS3-NS4A fusion proteins which can be used to identify inhibitors of this type, as well as enabling structural studies of the protease and protease: Inhibitor complexes. This sequence contains the alpha-helixo-Modified hepatitis C virus (HCV) NS3 protease comprising at least 1 substitution of a hydrophobic alpha-helix 0 amino acid to a hydrophilic amino acid, useful for screening inhibitors that may treat hepatitis CTSINGVLWTVYHGAGIRIIASPKGPVTQMYTNVDKDLVGWQAPQGSRSLTPCTCGSSDLY LVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVSTRGVAK 1 MKKKGSVVIVGRINLSGDTAYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVSTATQTFLA Hepatitis; NS3 protease; viral replication; chronic liver disease; liver failure; liver cancer; mutant; mutein. Goldfarb V; Score 985; DB 21; Pred. No. 3.2e-93; 0; Mismatches 6; Hepatitis C virus NS4A-NS3 fusion protease #2. Zhang Y, Ş (BRIM ) BRISTOL-MYERS SQUIBB CO. AAB15220 standard; protein; 195 Claim 23; Fig 13; 66pp; English AVDFIPVESLETIMRSP 197 AVDFIPVESLETTMRSP 197 Weinheimer S, Ouery Match Best Local Similarity 97.0%; Matches 191; Conservative 06-JAN-2000; 2000WO-US00345 99US-0115271 (first entry) 197 AA; Hepatitis C virus. Synthetic. WPI: 2000-465976, N-PSDB; AAA73330 WO200040707-A1. Wittekind M, 08-JAN-1999; 19-DEC-2000 13-JUL-2000 Sequence 61 121 121 181 181 AAB15220; ariant. **AAB1522**0 RESULT ö g ò g ö 셤 ò a

```
The present sequence is a mutated version of a fusion protein created using the Hepatitis C virus (HCV) NG3 and NS4A protease enzymes. These proteins are both essential for the replication of the virus, acting to cleave its replicative proteins from the polyprotein produced from the HCV genome. Inhibitors of the two proteins should be effective as antiviral treatments of HCV infection. This is useful as HCV can lead to chronic liver disease such as cirrhosis, liver failure and liver cancer. The present invention concerns a number of NS3 mutants and NS3-NS4A fusion proteins which can be used to identify inhibitors of this type, as well as enabling structural studies of the protease and proteins valued to the protease and protease. This sequence contains the alpha-helix0-1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               119 LVTRHADVIPVRRKGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVCTRGVAK 178
                                                                                                                                                                                                                                                                                                                                                                                                                                                                      9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      28
                                                        Modified hepatitis C virus (HCV) NS3 protease comprising at least 1 substitution of a hydrophobic alpha-helix 0 amino acid to a hydrophilic amino acid, useful for screening inhibitors that may treat hepatitis C
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       TSINGVLWTVYHGAGTRTIASPKGPVTQMYTNVDKDLVGWQAPQGSRSLTPCTCGSSDLY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    LVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVSTRGVAK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1 MKKKGSVVIVGRINLSGDTAYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVSTATQTFLA
                                                                                                                                                                                                                                                                                                                                                                                                                                          Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Hepatitis; NS3 protease; viral replication; chronic liver disease;
liver failure; liver cancer; mutant; mutein.
                                                                                                                                                                                                                                                                                                                                                                                                                                          ..
                                                                                                                                                                                                                                                                                                                                                                                                      Score 951; DB 21; Length 195;
Pred. No. 1e-89;
1; Mismatches 7; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Hepatitis C virus NS4A-NS3 fusion protease #8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ₹
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            (BRIM ) BRISTOL-MYERS SQUIBB CO.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      AAB15226 standard; protein; 197
                                                                                                                                    Claim 23; Fig 12; 66pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              AVDFIPVESLETIMRSP 195
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           181 AVDFIPVESLETIMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                       Query Match
93.5%;
Best Local Similarity 94.9%;
Matches 187; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             06-JAN-2000; 2000WO-US00345.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (first entry)
             WPI; 2000-465976/40.
                                                                                                                                                                                                                                                                                                                                                                           195 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Hepatitis C virus
                            N-PSDB; AAA73329
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             W0200040707-A1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          08-JAN-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   19-DEC-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             13-JUL-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Synthetic.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     AAB15226;
                                                                                                                                                                                                                                                                                                                                                                             Sequence
                                                                                                                                                                                                                                                                                                                                             variant
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           RESULT 7
AAB15226
ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                        ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                       9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    9
```

ä

Gaps

; 0

ë.

Length 197; Indels us-09-965-594-18.rag

n

```
28-JUL-1998;
28-NOV-1997;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Synthetic
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                complexes.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Seguence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           59
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                AAY24943;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Local
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            RESULT 9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ö
                                                                                                                                                                                                                                                    The present sequence is a mutated version of a fusion protein created using the Hepatitis C virus (HCV) NS3 and NS4A protease enzymes. These proteins are both essential for the replication of the virus, acting to cleave its replicative proteins from the polyprotein produced from the WCV genome. Inhibitors of the two proteins should be effective as antiviral treatments of HCV infection. This is useful as HCV can lead to chronic liver disease such as cirrhosis, liver failure and liver cancer. The present invention concerns a number of NS3 mutants and NS3-NS4A fusion proteins which can be used to identify inhibitors of this type, as well as enabling structural studies of the protease and protease. This sequence contains the alpha-helix0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   TSINGVLWTVYHGAGTRTIASPKGPVTQMYTNVDKDLVGWQAPQGSRSLTPCTCGSSDLY 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     LVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVSTRGVAK 180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1 MKKKGSVVIVGRINLSGDTAYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVSTATQTFLA 60
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Modified hepatitis C virus (HCV) NS3 protease comprising at least 1 substitution of a hydrophobic alpha-helix 0 amino acid to a hydrophilic amino acid, useful for screening inhibitors that may treat hepatitis C
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Hepatitis, NS3 protease, viral replication; chronic liver disease;
liver failure, liver cancer.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Length 197;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Indels
  Zhang Y, Goldfarb V;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Score 946; DB 21;
Pred. No. 3.3e-89;
0; Mismatches 11;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Hepatitis C virus NS4A-NS3 fusion protease #1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Ź
                                                                                                                                                                                                               Example 5; Fig 18; 66pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (BRIM ) BRISTOL-MYERS SQUIBB CO.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    AAB15212 standard; protein; 195
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     181 AVDFIPVESLETTMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            AVDFIPVESLETIMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Query Match
Best Local Similarity 94.4%;
Matches 186; Conservative (
  Weinheimer S,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              06-JAN-2000; 2000WO-US00345
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        (first entry)
                                        WPI; 2000-465976/40.
N-PSDB; AAA73335.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    197 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          W0200040707-A1.
  Wittekind M,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    19-DEC-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    13-JUL-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Synthetic
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              AAB15212;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Query Match
THE STATE OF THE S
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ద
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ð
```

```
TSINGVLWTVYHGAGTRTIASPKGPVTQMYTNVDKDLVGWQAPQGSRSLTPCTCGSSDLY 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              LVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVSTRGVAR 180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Modified hepatitis C virus (HCV) NS3 protease comprising at least 1 substitution of a hydrophobic alpha-helix 0 amino acid to a hydrophilic amino acid, useful for screening inhibitors that may treat hepatitis C
                                                                                                                                                                                                                                                                                                                                                          The present sequence is a fusion protein created using the Hepatitis C virus (HCV) NS3 and NS4A protease enzymes. These proteins are both essential for the replication of the virus, acting to cleave its replicative proteins from the polyprotein produced from the HCV genome. Inhibitors of the two proteins should be effective as antiviral treatments of HCV infection. This is useful as HCV can lead to chronic liver disease such as cirrhosis, liver failure and liver cancer. The present invention concerns a number of NS3 mutants and NS3-NS4A fusion proteins which can be used to identify inhibitors of this type, as well as enabling structural studies of the protease and protease:inhibitor
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1 MKKKGSVVIVGRINLSGDTAYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVSTATQTFLA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  HCV; hepatitis C virus; single chain recombinant complex; linker; NS3 protease; NS4A cofactor; NS4A-NS3 complex; serine protease; hydrophobic domain; covalent complex; detection; inhibitor.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       .;
;;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Score 912; DB 21; Length 195; Pred. No. 1e-85;
Goldfarb V;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    12;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1; Mismatches
Zhang Y,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Z
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           NS4A-NS3 complex SEQ ID NO:14.
                                                                                                                                                                                                                                                                                                          Example 2; Fig 10; 66pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     AAY24943 standard; Protein; 665
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   AVDFIPVESLETIMRSP 195
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      89.7%;
92.4%;
Wittekind M, Weinheimer S,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         98WO-0524528
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    98US-0094331
97US-0067315
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      AVDFIPVESLETTMRSP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             07-SEP-1999 (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Matches 182; Conservative
                                                       WPI; 2000-465976/40.
N-PSDB; AAA73328.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               195 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              W09928482-A2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      24-NOV-1998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            10-JUN-1999
```

ż

Yao

```
S2 CVNGVCWTVYHGAGSKTLAGPKGPITQMYTNVDQDLVGWQAPPGARSLTPCTCGSSDLYL 141
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   VIRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVSTRGVAKA 181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    142 VTRHADVIPVRRRGDSRGSLLSPRPVSYLKGSSGGPLLCPSGHAVGIFRAAVCTRGVAKA 201
                                                                                                                                                                                          The present invention describes a covalent hepatitis C virus (HCV) NS4A-NS3 complex comprising a central hydrophobic domain of native HCV NS4A peptide, a linker, and an HCV NS3 estine protease domain, where the hydrophobic domain of native HCV NS4A peptide is tethered by the linker to the amino terminus of the HCV NS3 protease domain. The present sequence represents a specifically claimed example of the above complex. The covalent NS4A-NS3 complexes are useful for structural determination and determination of mode of binding of HCV inhibitors by NMR spectroscopy. They can also be used for detecting inhibitors of the protease activity, the helicase activity and the ATPase activity of NS3. The covalent NS4A-NS3 complexes are more soluble, stable and active than the non-covalent protease-peptide complexes previously available.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          62 SINGVLWTVYHGAGTRIIASPKGPVTQMYTNVDKDLVGWQAPQGSRSLIPCTCGSSDLYL 121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      5 GSVVIVGRINLSGD---TAYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVSTATQTFLAT 61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          22 GSVVIVGRIILSGSGSITAYSQQTRGLLGCKKTSLTGRDKNQVEGEVQVVSTATQSFLAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          HCV; hepatitis C virus; single chain recombinant complex; linker; NS3 protease; NS4A cofactor; NS4A-NS3 complex; serine protease; hydrophobic domain; covalent complex; detection; inhibitor.
                                                                                                                                                                                                                                                                                                                                                                                                                                                             Query Match 86.7%; Score 881.5; DB 20; Length 665; Best Local Similarity 85.2%; Pred. No. 8e-82; Matches 167; Conservative 16; Mismatches 10; Indels 3;
                                                                                                                         New hepatitis C virus covalent complexes
                                                                                                                                                              Claim 6; Page 90-92; 211pp; English
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 AAY24947 standard; Protein; 665 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          HCV NS4A-NS3 complex SEQ ID NO:18.
                                                     PC,
                                                       Weber
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       VDFIPVESLETTMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          VDFVPVESMETTMRSP 217
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   98US-0094331.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          (first entry)
                                                       Taremi SS,
                   (SCHE ) SCHERING CORP
                                                                                         WPI; 1999-385385/32
                                                                                                                                                                                                                                                                                                                                                                                                                                 665 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        WO9928482-A2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               24-NOV-1998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   28-JUL-1998;
28-NOV-1997;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        07-SEP-1999
                                                       Malcolm BA,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            10-JUN-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Synthetic.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     AAY24947;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   122
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        202
                                                                                                                                                                                                                                                                                                                                                                                                                                   Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 RESULT 10
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  AAY24947
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          윱
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           셤
```

```
82 CVNGVCWTVYHGAGSKTLAGPKGPITQMYTNVDQDLVGWQAPPGARSLTPCTCGSSDLYL 141
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            62 SINGVLWTVYHGAGTRTIASPRGPVTOMYTNVDRDLVGWQAPQGSRSLTPCTCGSSDLYL 121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         122 VTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVSTRGVAKA 181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          NS4A.NS3 complex comprising a central hydrophobic domain of native NC4 NS4A peptide, a linker, and an HCV NS3 serine protease domain, where the hydrophobic domain of native HCV NS4A peptide is tethered by the linker to the amino terminus of the HCV NS4A peptide is tethered by the linker sequence represents a specifically claimed example of the above complex. The covalent NS4A-NS3 complexes are useful for structural determination and determination of mode of binding of HCV inhibitors by NNR spectroscopy. They can also be used for detecting inhibitors of the protease activity, the helicase activity and the Arpasa activity of NS3. The covalent NS4A-NS3 complexes are more soluble, stable and active than the non-covalent protease-peptide complexes previously available.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                22 GSVVIVGRIILSGSGSITAYSQQTRGLIGGKKTSLTGRDKNQVEGEVQVVSTATQSFLAT 81
                                                                                                                                                                                                                                                                                                                                                                                                                                                             Gaps
                                                                                                                                                           The present invention describes a covalent hepatitis C virus (HCV)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              HCV; hepatitis C virus; single chain recombinant complex; linker; NS3 protease; NS4A cofactor; NS4A-NS3 complex; serine protease; hydrophobic domain; covalent complex; detection; inhibitor.
                                                                                                                                                                                                                                                                                                                                                                                                                                e
N
                                                                                                                                                                                                                                                                                                                                                                                                Length 665;
                                                                                                                                                                                                                                                                                                                                                                                             / Match
Local Similarity 84.7%; Pred. No. 1.6e-81;
les 166; Conservative 17; Mismatches 10; Indels
                                ż
                                                                                             New hepatitis C virus covalent complexes
                                                                                                                             Claim 6; Page 100-102; 211pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ż
                                Weber PC,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 HCV NS4A-NS3 complex SEQ ID NO:12.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     AAY24941 standard; Protein; 665
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       VDFIPVESLETIMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      VDFVPVESMETTMRSP 217
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            98WO-US24528.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           98US-0094331.
97US-0067315.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  07-SEP-1999 (first entry)
                                Taremi SS,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            (SCHE ) SCHERING CORP
(SCHE ) SCHERING CORP
                                                            WPI; 1999-385385/32
                                                                                                                                                                                                                                                                                                                                                                   665 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            W09928482-A2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            24-NOV-199B;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            28-JUL-1998;
28-NOV-1997;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              10-JUN-1999
                                Malcolm BA,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Synthetic.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  AAY24941;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       182
                                                                                                                                                                                                                                                                                                                                                                   Sequence
                                                                                                                                                                                                                                                                                                                                                                                                  Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                   Best Loca
Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     RESULT 11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      AAY 24941
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      δý
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ŏ
```

8

3; Gaps

```
122 VTRHADVIPVRREGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVSTRGVAKA 181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   62 SINGVLWTVYHGAGIRTIASPKGPVTQMYTNVDKDLVGWQAPQGSRSLTPCTCGSSDLYL 121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          142 VTRHADVIPVRRRGDSRGSLLSPRPVSYLKGSSGGPLLCPSGHAVGIFRAAVCTRGVAKA 201
                                                                                                                                                                       NSAA PERIOR COMPISSION OF CENTRAL HYDROPHOBIC domain of native HCV NSAA peptide, a linker, and an HCV NS3 serine protease domain, where the hydrophobic domain of native HCV NSAA peptide is tethered by the linker to the amino terminus of the HCV NSAA peptide is tethered by the linker to the amino terminus of the HCV NS3 protease domain. The present sequence represents a specifically claimed example of the above complex. The covalent NSAA-NS3 complexes are useful for structural determination and determination of mode of binding of HCV inhibitors by Protease activity, the helicase activity and the ATPase activity of NS3. The covalent NSAA-NS3 complexes are more soluble, stable and active than the non-covalent protease-peptide complexes previously available.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   22 GSVVIVGRIILSGSGSITAYSQQTRGLLGCKITSLTGRDKNQYBGEVQVVSTATQSFLAT 81
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 5 GSVVIVGRINLSGD---TAYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVSTATQTFLAT 61
                                                                                                                                                            present invention describes a covalent hepatitis C virus (HCV)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             HCV; hepatitis C virus; single chain recombinant complex; linker;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              NS3 protease; NS4A cofactor; NS4A-NS3 complex; serine protease; hydrophobic domain; covalent complex; detection; inhibitor.
                                                                                                                                                                                                                                                                                                                                                                                                                                                             3;
                                                                                                                                                                                                                                                                                                                                                                                                                          86.3%; Score 877.5; DB 20; Length 665; 85.2%; Pred. No. 2.1e-81;
                                                                                                                                                                                                                                                                                                                                                                                                                                                             11; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                           15; Mismatches
                                                                                     New hepatitis C virus covalent complexes
                                                                                                                        Claim 6; Page 85-87; 211pp; English
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Ź
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          HCV NS4A-NS3 complex SEQ ID NO:13.
                <u>Б</u>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      AAY24942 standard; Protein; 665
                Weber
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           182 VDFIPVESLETTMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                202 VDFVPVESMETTMRSP 217
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             97US-0067315.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           98WO-US24528
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          98US-0094331
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         07-SEP-1999 (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                          Query Match
Best Local Similarity 85.2%
Matches 167; Conservative
                Malcolm BA, Taremi SS,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  (SCHE ) SCHERING CORP.
                                                    WPI; 1999-385385/32
                                                                                                                                                                                                                                                                                                                                                                                          665 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Hepatitis C virus.
Synthetic.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      W09928482-A2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             24-NOV-1998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              28-JUL-1998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 28-NOV-1997;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       10-JUN-1999
                                                                                                                                                                                                                                                                                                                                                                                            Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       AAY24942;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      RESULT 12
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         AAY24942
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ద
```

```
62 SINGVLWTVYHGAGTRTIASPKGPVTQMYTNVDKDLVGWQAPQGSRSLTPCTCGSSDLYL 121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      12 VTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVSTRGVAKA 181
                                                                                                                                          NS4A-NS3 complex comprising a central hydrophobic domain of native HCV NS3 aperide. Ws4A peptide. A a linker, and an HCV NS3 serine protease domain, where the hydrophobic domain of native HCV NS4A peptide is tethered by the linker to the amino terminus of the HCV NS3 protease domain. The present sequence represents a specifically claimed example of the above complex. The covalent NS4A-NS3 complexes are useful for structural determination and determination of mode of binding of HCV inhibitors by NNR spectroscopy. They can also be used for detecting inhibitors of the protease activity, the helicase activity and the Arpase activity of NS3 The covalent NS4A-NS3 complexes are more soluble, stable and active than
                                                                                                                                                                                                                                                                                                                                                                                                                                                         22 GSVVIVGRIILSGSGSITAYSQQTRGLLGCIKTSLTGRDKNQVBGEVQVVSTATQSFLAT
                                                                                                                                                                                                                                                                                                                                                                                                                                        5 GSVVIVGRINLSGD---TAYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVSTATQTFLAT
                                                                                                                                                                                                                                                                                                                                                                                                         Gaps
                                                                                                                            present invention describes a covalent hepatitis C virus (HCV)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           HCV; hepatitis C virus; single chain recombinant complex; linker; NS3 protease; NS4A cofactor; NS4A-NS3 complex; serine protease; hydrophobic domain; covalent complex; detection; inhibitor.
                                                                                                                                                                                                                                                                                                          the non-covalent protease-peptide complexes previously available.
                                                                                                                                                                                                                                                                                                                                                                                                      3;
                                                                                                                                                                                                                                                                                                                                                                      86.3%; Score 877.5; DB 20; Length 665; 85.2%; Pred. No. 2.1e-81; ative 15; Mismatches 11; Indels 3;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Yao N;
ż
                                                             New hepatitis C virus covalent complexes
                                                                                             Claim 6; Page 88-90; 211pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Weber PC,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    $
P,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              HCV NS4A-NS3 complex SEQ ID NO:4.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  AAX17880 standard; Protein; 216
Weber
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    VDFIPVESLETTMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   UDEVPVESMETTMRSP 217
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              98WO-US24528
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           97US-0067315
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             98US-0094331
                                                                                                                                                                                                                                                                                                                                                                                     Local Similarity 85.2 les 167; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Malcolm BA, Taremi SS,
Malcolm BA, Taremi SS,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (SCHE ) SCHERING CORP
                              WPI; 1999-385385/32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Hepatitis C virus
                                                                                                                                                                                                                                                                                                                                         665 A.A.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          28-JUL-1998;
28-NOV-1997;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              W09928482-A2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           24 - NOV - 1998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                07-SEP-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Synthetic.
                                                                                                                                                                                                                                                                                                                                          Seguence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 AAY17880;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    182
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  202
                                                                                                                                                                                                                                                                                                                                                                       Query Match
                                                                                                                                                                                                                                                                                                                                                                                                        Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    RESULT 13
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 AAY17880
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
```

```
62 SINGVLWTVYHGAGTRTIASPKGPVTQMYTNVDKDLVGWQAPQGSRSLTPCTCGSSDLYL 121
                                                                                                          The present invention describes a covalent hepatitis C virus (HCV) NS4A-NS3 complex comprising a central hydrophobic domain of native HCV NS4A-NS3 complex. Infact, and an HCV NS3 Serine procease domain, where the hydrophobic domain of native HCV NS4A peptide is tethered by the linker to the amino terminus of the HCV NS3 protease domain. The present sequence represents a specifically claimed example of the above complex. The covalent NS4A-NS3 complexes are useful for structural determination and determination of mode of binding of HCV inhibitors by NNR spectroscopy. They can also be used for detecting inhibitors of the protease activity, the helicase activity and the Arpase activity of NS3. The covalent NS4A-NS3 complexes are more soluble, stable and active than the non-covalent protease-peptide complexes previously available.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           VTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVSTRGVAKA 181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            5 GSVVIVGRINLSGD -- - TAYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVSTATQTFLAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                         GSVVIVGRIILSGSGSITAYSQQTRGLLGCKKTSLTGKDKNQVEGEVQVVSTATQSFLAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   HCV; hepatitis C virus; single chain recombinant complex; linker; NS3 protease; NS4A cofactor; NS4A-NS3 complex; serine protease; hydrophobic domain; covalent complex; detection; inhibitor.
                                                                                                                                                                                                                                                                                                                                                            Query Match 86.0%; Score 874.5; DB 20; Best Local Similarity 85.1%; Pred. No. 8.8e-82; Matches 166; Conservative 16; Mismatches 10;
                                                 New hepatitis C virus covalent complexes
                                                                                Claim 6; Page 76-77; 211pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Weber PC,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         HCV NS4A-NS3 complex SEQ ID NO:16
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          VDFVPVESMETTMRS 216
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          VDFIPVESLETTMRS 196
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       AAY24945 standard; Protein;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        98WO-US24528
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       98US-0094331
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          97US-0067315
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Taremi SS,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (SCHE ) SCHERING CORP
                  WPI; 1999-385385/32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Hepatitis C virus Synthetic.
                                                                                                                                                                                                                                                                                                                              216 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        WO9928482-A2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      24-NOV-1998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       28-JUL-1998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          28-NOV-1997;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        07-SEP-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       10-JUN-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          BA,
                                                                                                                                                                                                                                                                                                                              Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                          22
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      202
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        AAY24945;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Malcolm
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          RESULT 14
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           AAY24945
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        셤
q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      a
                                                                                                                                                                                                                                                                                                                                                                                                                           ö
```

Yao N;

Length 216; Indels

```
82 CVNGVCWTVYHGAGSKTLAGPKGPITQMYINVDQDLVGWQAPPGARSLTPCTCGSSDLYL 141
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    122 VTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVSTRGVAKA 181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SINGVLWTVYHGAGTRIIASPKGPVTQMYTNVDKDLVGWQAPQGSRSLTPCTCGSSDLYL 121
                                                                                   The present invention describes a covalent hepatitis C virus (HCV) NS4A-NS3 complex comprising a central hydrophobic domain of native HCV NS4 serine protease domain, where the hydrophobic domain of native HCV NS4A peptide is tethered by the linker to the amino terminus of the HCV NS4A peptide is tethered by the linker sequence represents a specifically claimed example of the above complex. The covalent NS4A NS3 complexes are useful for structural determination and determination of mode of binding of HCV inhibitors by NMR spectroscopy. They can also be used for detecting inhibitors of the protease activity and the ATPase activity of NS3 The covalent NS4A-NS3 complexes are more soluble, stable and active than
                                                                                                                                                                                                                                                                                                                                                                                                           81
                                                                                                                                                                                                                                                                                                                                                                              5 GSVVIVGRINLSGD---TAYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVSTATGTFLAT 61
                                                                                                                                                                                                                                                                                                                                                                                                 22 GSVVIVGRIILSGSGSITAYSQQTRGLLGCKITSLTGRDKNQVEGEVQVVSTATQSFLAT
                                                                                                                                                                                                                                                                                                                                                    Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             HCV; hepatitis C virus; single chain recombinant complex; linker; NS3 protease; NS4A cofactor; NS4A-NS3 complex; serine protease; hydrophobic domain; covalent complex; detection; inhibitor.
                                                                                                                                                                                                                                                            the non-covalent protease-peptide complexes previously available
                                                                                                                                                                                                                                                                                                                                                   3,
                                                                                                                                                                                                                                                                                                                      DB 20; Length 665;
                                                                                                                                                                                                                                                                                                                                                   Indels
                                                                                                                                                                                                                                                                                                                                                 11;
                                                                                                                                                                                                                                                                                                                 86.0%; Score 874.5; DB 2
84.7%; Pred. No. 4.2e-81;
Live 16; Mismatches 11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Yao N;
                              New hepatitis C virus covalent complexes
                                                         Page 95-97; 211pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Z
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Weber PC,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  HCV NS4A-NS3 complex SEQ ID NO:17.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      AAY24946 standard; Protein; 665
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               UDEVPVESMETTMRSP 217
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 182 VDFIPVESLETIMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                98US-0094331.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   98WO-US24528
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (first entry)
                                                                                                                                                                                                                                                                                                                                 Local Similarity 84.79 nes 166; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Taremi SS,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              (SCHE ) SCHERING CORP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             WPI; 1999-385385/32
WPI; 1999-385385/32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Hepatitis C virus.
                                                                                                                                                                                                                                                                                       665 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   24-NOV-1998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       W09928482-A2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                28-JUL-1998;
28-NOV-1997;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   07-SEP-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    10-JUN-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Malcolm BA,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Synthetic
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     AAY24946;
                                                                                                                                                                                                                                                                                       Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                          62
                                                           Claim 6;
                                                                                                                                                                                                                                                                                                                      Query Match
                                                                                                                                                                                                                                                                                                                                      Best Loca
Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           RESULT 15
AAY24946
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      õ
                                                                                                                                                                                                                                                                                                                                                                                                        qq
                                                                                                                                                                                                                                                                                                                                                                                                                                      ó
                                                                                                                                                                                                                                                                                                                                                                                                                                                                     요
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            8
                                                                                                                                                                                                                                                                                                                                                                                                                          81
                                                                                                                                                                                                                                                                                                                                                                 Gaps
                                                                                                                                                                                                                                                                                                                                                                3;
```

```
NS4A-NS3 complex comprising a central hydrophobic domain of native HCV NS4A-NS3 complex comprising a central hydrophobic domain of native HCV NS3 serine protease domain, where the hydrophobic domain of native HCV NS4A peptide is tethered by the linker to the amino terminus of the HCV NS3 protease domain. The present sequence represents a specifically claimed example of the above complex. The covalent NS4A-NS3 complexes are useful for structural determination and determination of mode of binding of HCV inhibitors by NMR spectroscopy. They can also be used for detecting inhibitors of the protease activity, the helicase activity and the Arpase activity of NS3. The covalent NS4A-NS3 complexes are more soluble, stable and active than the non-covalent protease-peptide complexes previously available.
                                                                                                                                           present invention describes a covalent hepatitis C virus (HCV)
                       New hepatitis C virus covalent complexes
                                                                                Claim 6; Page 97-99; 211pp; English.
```

665 AA; Sequence

3; Gaps Query Match
86.0%; Score 874.5; DB 20; Length 665;
Best Local Similarity 84.7%; Pred. No. 4.2e-81;
Matches 166; Conservative 16; Mismatches 11; Indels 3; δ

62 SINGVLWTVYHGAGTRIIASPKGPVTQMYTNVDKDLVGWQAPQGSRSLTPCTCGSSDLYL 121

q Š

q õ

g ò VDFIPVESLETTMRSP 197 |||:|||:||||||||||| VDFVPVESMETTMRSP 217 182

Search completed: August 30, 2003, 19:12:24 Job time : 45.6227 secs

```
GenCore version 5.1.6 Copyright (c) 1993 - 2003 Compugen Ltd.
```

Control of the second s	On process: Process Search, using Sw Modes.  Run on: August 30, 2003, 19:02:22 ; Search time 16.2134 Seconds	<pre>(without alignments) 1168.492 Million cell updates/sec</pre>		Perfect Score: 101/ Sequence: 1 MKKKGSVVIVGRINLSGDTAVAKAVDFIPVESLETTMRSP 197
Š	Run R		Titl	Sequ

283308 segs, 96168682 residues

Searched:

BLOSUM62 Gapop 10.0 , Gapext 0.5

Scoring table:

283308 Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

PIR\_76:\*
1: pir1:\*
2: pir2:\*
3: pir3:\*
4: pir4:\* Database :

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

hypothetical prote transferrin precur	irregular chiasm C hypothetical prote	transferrin-bindin genome polyprotein	probable exported ferredoxin-nitrite	transierrin-bindin surface-associated polyketide synthas	proteinase (EC 3.4 mosA protein - Rhi	htrA-like serine p hypothetical prote	
S57895 S33761	A49448 T04533	JN0821 GNWVTB	S16603	C81832 S61441 T17410	PRLJHD B53308	H97199 G90654	G85505
77	00	<b>~</b> ~ (	<b>7</b> ~ (	900	~ ~	~ ~	~
361	764	3412	400	3739	323	433	452
7.7	7.7			0.7.0	7.5	7.5	7.5
78.5	78.5	78	C.//	77 77	76.5 76.5	76.5 76.5	76.5
31	333	₩ ₩ ₩	3.0	0 E 4	41	4 4 6 4	5 4 5

## ALIGNMENTS

RESULT 1 GNWVC3 genome polyprotein - hepatitis C virus (strain HCV-1) genome polyprotein - hepatitis C virus (strain HCV-1) N:Contains: capsid protein C; envelope protein M; hepacivirin (EC 3.4.21.98) (nonstrainstein NS4s; nonstructural protein NS4s; nonstructural protein NS5 C; Secies: hepatitis C virus C; Species: hepatitis C virus C; Date: 30.5ep-1992 #sequence_revision 30-Sep-1992 #text_change 19-Jan-2001 C; Accession: A39166; PO0404 R; Choo, O.L.; Richman, K.H.; Han, J.H.; Berger, K.; Lee, C.; Dong, C.; Gallegos, C.; Proc, Natl. Acad. Sci. U.S.A. 88, 2451-2455, 1991 A; Title: Genetic organization and diversity of the hepatitis C virus. A; Reference number: A39166; MUD: 91172826; PMID: 1848704 A; Accession: A39166
A; Residues: 1.3011 <choa. 1.3011="" 1131-1141,="" 1316939="" 1992.="" 73,="" 92268871;="" <choa.="" a="" a;="" aaa45676.1;="" accession:="" analysis="" and="" b.;="" c="" chan.="" cross="" dow,="" e.;="" e.c.;="" f.;="" follett,="" g329873;="" g329874="" gb:="" gen.="" hepatitis="" holmes,="" its="" j.="" j.f.;="" m6.2321;="" mccomish,="" muid:="" new="" nid:="" number:="" of="" peutherer,="" phylogenetic="" pid:="" pidn:="" pnid:="" pq0393;="" pq0403<="" r;="" reference="" references:="" relationship="" residues:="" s.w.;="" td="" title:="" type="" virol.="" virus="" yap,=""></choa.>
A; Molecule type: genomic RNA A; Residues: 1577-1633 <cha <ned="" a;="" actaus:="" c="" c;="" cross="" ddbj-10128="" e-b16="" e-b17="" experimental="" genome="" genomic="" hepatitis="" hydrolase;="" isolates="" molecule="" nonstructural="" polyprotein="" predicted="" preliminary="" protein="" r;="" references:="" reperimental="" rna="" source:="" status="" status:="" superfamily:="" type:="" virus=""> F; 116-191/Product: nonstructural protein NS1 *status predicted <ns2> F; 1007-1615/Product: hepacivirin *status predicted <ns2></ns2></ns2></cha>
F;1220-1237/Region: nucleotide-binding motif A (P-10op) F;1310-1317/Region: nucleotide-binding motif B F;1316-1319/Region: nucleotide-binding motif B F;1316-1319/Region: DEXH motif F;1616-1862/Product: nonstructural protein NS40 *status predicted <nad>F;1863-2013/Product: nonstructural protein NS40 *status predicted <nad>F;2014-3011/Product: nonstructural protein NS5 *status predicted <nad>F;2014-3011/Product: nonstructural protein NS5 *status predicted <nad>F;196,209,234,305,325,417,423,430,448,476,532,540,556,576,623,645,1213,1255,2041,20; Query Match 84 0%; Score 854.5; DB 1; Length 3011; Best Local Similarity 82.8%; Pred. No. 2.7e-69; Matches 169; Conservative 9; Mismatches 17; Indels 9; Gaps 1;</nad></nad></nad></nad>

----LSGDTAYAQQTRCEBGCQETSQTGRDKNOVEGEVQIVST 53

3 KKGSVVIVGRIN-----

à

```
A; Molecule type: genomic RNA
A; Residues: 1-3011 <INC>
A; Residues: 1-3011 <INC>
A; Cross-references: GB: MG7463; NID: g329737; PIDN: AAA45534.1; PID: g329738
A; Cross-references: GB: MG7466; D. B.; Sugitani, M.; Nasoff, M.; Prince, A.W.
Proc. Natl. Acad. Sci. U.S.A. 88, 10292-10296, 1991
A; Title: Genomic structure of the human prototype strain H of hepatitis C virus: comp. A; Reference number: A1546; MUID: 92052256; PMID: 1658800
A; Contents: annotation
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       A) Note: neither amino acid nor nucleotide sequence is given
C; Superfamily: hepatitis C virus genome polyprotein
C; Superfamily: hepatitis C virus genome polyprotein
C; Superfamily: hepatitis C virus genome polyprotein; divorated representation of the protein 
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    File16-1862/Product: nonstructural protein NS4a *status predicted <N4A>
File16-1862-2013/Product: nonstructural protein NS4b *status predicted <NNB>
File14-3011/Product: nonstructural protein NS5 bstatus predicted <NS5>
File196,209,234,305,325,417,423,430,448,476,532,540,556,576,623,645,1213,1255,2041,2246
                                                                                                                                                                      genome polyprotein - hepatitis C virus (strain H)
N:Contains: capsid protein C; envelope protein M; hepacivirin (EC 3.4.21.98) (nonstru-
protein NS4a; nonstructural protein NS4b; nonstructural protein NS5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            A; Description: Genomic structure of the human prototype strain H of hepatitis C virus
                                                                                                                                                                                                                                                                                                                               A; Note: host Homo sapiens (man)
C; Date: 31-Dec-1992 #sequence_revision 31-Dec-1992 #text_change 19-Jan-2001
C; Accession: A36814; A41546
R; Inchauspe, G.; Zebedee, S.; Lee, D.H.; Sugitani, M.; Nasoff, M.; Prince, A.M. submitted to GenBank, July 1992
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1185 CTRGVAKAVDFIPVENLETTMRSP 1208
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      STRGVAKAVDFIPVESLETTMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    F;1316-1319/Region: DEXH motif
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Best Local Similarity 81.9
Matches 167; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Species: hepatitis C virus
                                                                                                                                                                                                                                                                                                    C; Species: hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                A; Reference number: A36814
A; Accession: A36814
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ç
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         A. Experimental source: isolate HG-JI
G. Superfamily: hepatitis G virus genome polyprotein
G. Superfamily: hepatitis G virus genome polyprotein
G. Superfamily: hepatitis G virus genome polyprotein
F. 2-115/Product: capsid protein G % status predicted <CPC>
F. 116-191/Product: capsid protein M % status predicted <EPM>
F. 116-191/Product: major envelope protein E % status predicted <MEE>
F. 390-729/Product: major envelope protein E % status predicted <MEE>
F. 730-1006/Product: nonstructural protein NS2 % status predicted <NS2>
F. 730-101/F. Product: nonstructural protein NS2 % status predicted <NS2>
F. 1230-1231/Region: nucleotide-binding motif A (P-loop)
F. 1312-1311/Region: nucleotide-binding motif B
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              genome polyprotein - hepatitis C virus
N;Contains: capsid protein C; envelope protein M; hepacivirin (EC 3.4.21.98) (nonstructu protein NS4a; nonstructural protein NS5 C;Schezis, hepatitis C virus
C;Date: 19-May-2000 *sequence_revision 19-May-2000 *text_change 19-Jan-2001
C;Accession: $40770; PC1285
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           A;Cross-references: EMBL:D10749; NID:g221586; PIDN:BAA01582.1; PID:g221587
R;Okamoto, H.; Okada, S.; Sugiyama, Y.; Yotsumoto, S.; Tanaka, T.; Yoshizawa, H.; Tsuda, Jpn. J. Exp. Med. 60, 167-177, 1990
A;Title: The S'-terminal sequence of the hepatitis C virus genome.
A;Reference number: PC1284; MUID:91013116; PMID:2170712
A;Accession: PC1285
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1005 RKGREILLGPADGMVSKGWRLLAPITAYAQQTRGLLGCIITSLTGRDKNQVEGEVQIVST 1064
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       113
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     173
                                                                                                                                                                                                                                                                                                    1005 RRGREILLGPADGMVSKGWRLLAPITAYAQQTRGLLGCIITSLTGRDKNQVEGEVQIVST 1064
                                                                                          54 ATQTFLATSINGVLWTVYHGAGTRTIASPKGPVTQMYTNVDKDLVGWQAPQGSRSLTPCT 113
                                                                                                                                                                                                                                                                114 CGSSDLYLVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAV 173
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              54 ATQTFLATSINGVLWTVYHGAGTRTIASPKGPVTQMYTNVDKDLVGWQAPQGSRSLTPCT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   3 KKGSVVIVGRIN-----LSGDTAYAQQTRGEEGCQETSQTGRDRNQVEGEVQIVST
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     114 CGSSDLYLVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAV
                                                                                                                                       A;Cross-references: GB:D00831; NID:9221511; PIDN:BAA00705.1; PID:9221512
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      F;1616-1862/Product: nonstructural protein NS4a #status predicted <N4A>
F;1863-2013/Product: nonstructural protein NS4b #status predicted <N4B>
F;2014-3011/Product: nonstructural protein NS5 #status predicted <NS5>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Length 3011;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Score 853.5; DB 1;
Pred. No. 3.4e-69;
8; Mismatches 18;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               R;Okamoto, H. submitted to the EMBL Data Library, March 1992 A;Reference number: $40770 A;Accession: $40770
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1185 CIRGVAKAVDFIPVESLETIMRSP 1208
                                                                                                                                                                                                                                                                                                                                                                                                                                                                               STRGVAKAVDFIPVESLETTMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              .;
8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 83.9%;
82.8%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             F;1316-1319/Region: DEXH motif
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    A; Molecule type: genomic RNA A; Residues: 1-513 < OK2>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            A; Molecule type: genomic RNA
A; Residues: 1-3011 < OKA>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Query Match
Best Local Similarity
Matches 169; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           174
                                                                                                                                                                                                                                                                                                                                                                                                                                  174
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ô
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            셤
                                                                                                                                                                      g
                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                  ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
    셤
                                                                                          ò
```

```
genome polyprotein - hepatitis C virus (strain Taiwan)
N;Contains: capsid protein C; envelope protein M; hepacivirin (EC 3.4.21.98) (nonstru
protein NS4a; nonstructural protein NS4b; nonstructural protein NS5
                                                                                                                                                                                                                     1005 RRGQEILLGPADGMVSKGWRLLAPITAYAQQTRGLLGCIITSLIGRDKNQVEGEVQIVST 1064
                                                                                                                                                                                                                                                                                                                                                                   1065 ATQTFLATCINGVCWIVYHGAGTRTIASPKGPVIQTYINVDQDLVGWPAPQGSRSLTPCT 1124
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                113
                                                                                                                                                                                                                                                                                                                                                                                                                                       114 CGSSDLYLVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAV 173
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   A;Note: host Homo sapiens (man)
C;Date: 31-Dec-1992 #sequence_revision 31-Dec-1992 #text_change 19-Jan-2001
                                                                                                                                                     3 KKGSVVIVGRIN-----LSGDTAYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVST
                                                                                                                                                                                                                                                                                                54 ATQTFLATSINGVLWTVYHGAGTRTIASPKGPVTQMYTNVDKDLVGWQAPQGSRSLTPCT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           C; Accession: A40244
R;Chen, P.J.; Lin, M.H.; Tai, K.F.; Liu, P.C.; Lin, C.J.; Chen, D.S.
Virology 188, 102-113, 1992
       Length 3011;
                                                                                                                                                                                                                                                                                                                                     Indels
           DB 1;
83.4%; Score 848.5; DB 1
81.9%; Pred. No. 9.7e-69;
tive 10; Mismatches 18
```

```
F;1616-1862/Product:
F;1863-2013/Product:
F;2014-3010/Product:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1065
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1125
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   114
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       genome
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ద
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                               a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       셤
                                                                                                                                                                                                                                                                                  õ
                                                                                                                                                                                                                                                                                                                                                                                                                       à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     A: Experimental source: HCV-JT
A: Note: sequence extracted from NCBI backbone (NCBIN:106206, NCBIP:106207)
C: Superimental source: HCV-JT
A: Note: sequence extracted from NCBI backbone (NCBIN:106206, NCBIP:106207)
C: Superimanly: hepatitis C virus genome polyprotein
F: 2-115/Product: capsid protein C *status predicted <PED>
F: 116-191/Product: envelope protein E *status predicted <NSI>
F: 390-729/Product: nonstructural protein NSI *status predicted <NSI>
F: 300-1066/Product: nonstructural protein NSI *status predicted <NSI>
F: 1230-1237/Region: nucleotide-binding motif A (P-loop)
F: 1312-1317/Region: nucleotide-binding motif B
                                                                                                                                                                                                                                      C.Keywords: ATP; capsid protein; envelope protein; glycoprotein; hydrolase: nonstructural product: capsid protein C *status predicted <CPC>
F;116-191/Product: capsid protein C *status predicted <CPC>
F;16-191/Product: envelope protein M *status predicted <CPC>
F;192-189/Product: major envelope protein Status predicted <MRE>
F;390-729/Product: nonstructural protein NSI *status predicted <NSI>
F;730-1006/Product: nonstructural protein NSI *status predicted <NSI>
F;1007-1615/Product: hepacivirin *status predicted <NSI>
F;1230-1237/Region: nucleotide-binding motif A (P-loop)
F;1312-1317/Region: DEXH motif
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           genome polyprotein - hepatitis C virus (strain JT)

N;Contains: capsid protein C; envelope protein M; hepacivirin (EC 3.4.21.98) (nonstructu protein NS4a; nonstructural protein NS4a; nonstructural protein NS5C;Species: hepatitis C virus

C;Species: hepatitis C virus

C;Date: 19-May-2000 #sequence_revision 19-May-2000 #text_change 19-Jan-2001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             F;1616-1862/Product: nonstructural protein NS4a #status predicted <N4A>
F;1863-2013/Product: nonstructural protein NS4b #status predicted <N4B>
F;2014-3010/Product: nonstructural protein NS5 #status predicted <NS5>
F;2014-3010/Product: nonstructural protein NS5 #status predicted <NS5>
F;196,209,233,214,250,305,325,417,423,430,448,532,540,556,576,623,645,1213,1255,2041,207
A:Title: The Taiwanese hepatitis C virus genome: sequence determination and mapping the A:Reference number: A40244; MUID:92230206; PMID:1314449
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Ritanaka, T.; Kato, N.; Nakagawa, M.; Ootsuyama, Y.; Cho, M.J.; Nakazawa, T.; Hijikata, Virus Res. 23, 39-53, 1992
A; Title: Molecular cloning of hepatitis C virus genome from a single Japanese carrier: A; Reference number: A45573; MUID:92295714; PMID:1318627
A; Raccession: A45573
A; Molecule type: DNA
A; Molecule type: DNA
A; Residues: 1-3010 CTAN>
A; Cross-references: GB:D11168; GB:D01171; NID:9221612; PIDN:BAA01943.1; PID:9221613
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1005 RRGREILLGPADSLEGRGWRLLAPITAYAQQTRGLFGCIITSLTGRDKNQVEGEVQVVST 1064
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CGSSDLYLVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   3 KKGSVVIVGRIN------LSGDTAYAQOTRGEEGCQETSQTGRDKNQVEGEVQIVST
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ATQTFLATSINGVLWTVYHGAGTRTIASPKGPVTQMYTNVDKDLVGWQAPQGSRSLTPCT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Length 3010;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          DB 1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       82.4%; Score 837.5; DB 178.4%; Pred. No. 9.8e-68;
                                                                                                                                                                         A;Cross-references: GB:M84754
C;Superfamily: hepatitis C virus genome polyprotein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           18; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CTRGVAKAVDFVPVESMETTMRSP 1208
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    174 STRGVAKAVDFIPVESLETTMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    F;1312-1317/Region: nucleotide F;1316-1319/Region: DEXH motif
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Query Match 82.4's
Best Local Similarity 78.4'
Matches 160; Conservative
                                                                                                        A; Molecule type: genomic RNA
A; Residues: 1-3010 <CHE>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         C; Accession: A45573
                                                                           A; Accession: A40244
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ç
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ð
```

```
A/KESTORS-references: EMEL:M58335; NID:9329770; PIDN:AAA72945.1; PID:9329771
C; Superfamily: hepatitis C virus genome polyprotein
C; Superfamily: hepatitis C virus genome polyprotein
C; Superfamily: hepatitis C virus genome polyprotein
C; Keywords: ATP; capsid protein; envelope protein; glycoprotein; hydrolase; nonstruc
C; Keywords: ATP; capsid protein B * Status predicted < CEPC>
F: 2-115.Product: envelope protein M * Status predicted < CEPN>
F: 192-389/Product: major envelope protein B * Status predicted < MES>
F: 190-729/Product: nonstructural protein NS1 * Status predicted < NS2>
F: 130-1006/Product: nonstructural protein NS2 * Status predicted < NS2>
F: 130-1217/Region: nucleotide-binding motif A (P-loop)
F: 1312-1317/Region: nucleotide-binding motif B
F: 1316-1319/Region: nucleotide-binding motif B
F: 1316-1319/Region: nucleotide-binding motif B
F: 1316-1319/Region: nonstructural protein NS4* * Status predicted < N4A>
F: 2014/Product: nonstructural protein NS5* * Status predicted < NS5>
F: 2014-3013/Product: nonstructural protein NS5* * Status predicted < NSS>
F: 2014-3010/Product: nonstructural protein NS5* * Status predicted < NSS>
F: 196, 209, 234, 250, 305, 325, 417, 423, 430, 448, 532, 540, 556, 576, 623, 645, 1213, 1255, 2041, 207
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    N.Contains: capsid protein C; envelope protein M; hepacivirin (EC 3.4.21.98) (nonstranctein NS4s; nonstructural protein NS4b; nonstructural protein NS5 (Species: hepatitis C virus (Species: hepatitis C virus (Species: har-1992 *sequence_revision 31-Mar-1992 *text_change 19-Jan-2001 (Species: 31-Mar-1992 *sequence_revision 31-Mar-1992 *text_change 19-Jan-2001 (Species: 31-Mar-1992 *sequence_revision 31-Mar-1992 *text_change 19-Jan-2001 (Species: 31-Mar) (Species: 100-1113, 1991)

3. Virol. 65, 1105-1113, 1991

3. Title: Structure and organization of the hepatitis C virus genome isolated from hu A; Reference number: A38465; MuID:91140698; PMID:1847440
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        RRGKEILLGPADSLEGRGLRLLAPITAYSQQTRGLLGCIITSLTGRDKNQVEGEVQVVST 1064
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ATQSFLATCVNGVCWTVYHGAGSKTLAAPKGPITQMYTNVDQDLVGWPKPPGARSLTPCT 1124
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CGSSDLYLVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAV 173
                                                                                                                                                                                                                                                                                                                                                                                                                     1005 RRGREILLGPADSIEGGGWRLLAPITAYAQOTRGLLGCIVTSLTGRDKNQVEGEVQVVST 1064
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         114 CGSSDLYLVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAV 173
                                                                                                                                                                                                                                                                                                                                      53
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                54 ATQTFLATSINGVLWTVYRGAGIRTIASPKGPVTQMYTNVDKDLVGWQAPQGSRSLTPCT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          3 KKGSVVIVGRIN------LSGDTAYAQOTRGEEGCQETSQTGRDKNQVEGEVQIVST
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   54 ATQTFLATSINGVLWTVYHGAGTRTIASPKGPVTQMYTNVDKDLVGWQAPQGSRSLTPCT
                                                                                                                                                                                                                                                                                                                                      3 KKGSVVIVGRIN-----LSGDTAYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVST
nonstructural protein NS4a #status predicted <N4A> nonstructural protein NS4b #status predicted <N4B> nonstructural protein NS5 #status predicted <NS5>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       6
                                                                                                                                                                                                                                                   6
                                                                                                                                                                 Length 3010;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       18; Indels
                                                                                                                                                                                                                                                   Indels
                                                                                                                                                    ; Score 827.5; DB 1;
; Pred. No. 8e-67;
20; Mismatches 18;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                DB 1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ; Score 823.5; DB 1
; Pred. No. 1.9e-66;
21; Mismatches 18
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1185 CTRGVAKAVDFIPVESMETTMRSP 1208
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              174 STRGVAKAVDFIPVESLETTMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         polyprotein - hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     81.0%;
76.5%;
                                                                                                                                                                 81.48;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Best Local Similarity ....
Matches 156; Conservative
                                                                                                                                                            Query Match
Best Local Similarity 77.0°
Matches 157; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      A; Molecule type: genomic RNA
A; Residues: 1-3010 <TAK>
```

```
As note: this sequence is inconsistent with the nucleotide translation
A;Note: this sequence is inconsistent with the nucleotide translation
A;Note: the authors translated the codon AGG for residue 43 as Pro, TGG for residue 3
as Trp, and Trc for residue 771 as Ser
A;Note: sequence extracted from NCBI backbone (NCBIN:121747, NCBIP:121748)
C;Superfamily: hepatitis C virus genome polyprotein
C;Superfamily: hepatitis C virus genome polyprotein
C;Superfamily: hepatitis C virus genome polyprotein
F;2-115/Product: capsid protein P * status predicted <PDM>F;116-191/Product: capsid protein M * status predicted <PDM>F;130-729/Product: nonstructural protein NSI * status predicted <NSI>F;730-1006/Product: nonstructural protein NSI * status predicted <NSI>F;130-1006/Product: nonstructural protein NSI * status predicted <NSI>F;130-1006/Product: nonstructural protein NSI * status predicted <NSI>F;1310-1237/Region: nucleotide-binding motif A (P-loop)
F;1312-1317/Region: DEXH motif
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             C; Species: hepatitis C virus
C; Species: hepatitis C virus
C; Stele: 19-May-2000 #sequence_revision 19-May-2000 #text_change 19-Jan-2001
C; Accession: JC5620
R; Chamberlain, R.W.; Adams, N.J.; Taylor, L.A.; Simmonds, P.; Elliott, R.W.
Biochem. Biochem. Biochem. 205, 44-49, 1997
A; Title: The complete coding sequence of hepatitis C virus genotype 5a, the predomina A; Reference number: JC5620; MUID: 97366593; PMID: 9223423
                                C;Date: 19-May-2000 *sequence_revision 19-May-2000 *text_change 23-Mar-2001 C;Accession: S18030, S33570, A48332; S18029 **
C;Accession: S18030, S33570, A48332; S18029 **
Submitted to the EMBL Data Library, September 1991 A;Bescription: A whole genome of hepatitis C virus CDNA was isolated from a single pa A;Reference number: S18028 A;Reference number: S18038
                                                                                                                                                                                                                                                                                                  A; Molecule type: genomic RNA
A; Residues: 1-3010 <HON>
A; Residues: 1-3010 <HON>
A; Residues: BMBL: X61596; NID: 959478; PIDN: CAA43793.1; PID: 959479
A; Experimental source: 1solate JK1 from an individual
B; Honda, M.; Kaneko, S.; Unoura, H.; Kobayashi, K.; Murakami, S.
Arch. Virol. 128, 163-169, 1993
A; Title: Sequence analysis of putative structural regions of hepatitis C virus isolat
A; Reference number: A48332; MUID: 93119270; PMID: 8380322
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    N:Contains: capsid protein C; envelope protein M; hepacivirin (EC 3.4.21.98) (nonstruprotein NS4a; nonstructural protein NS4b; nonstructural protein NS5 C; Species: hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          F:1616-1862/Product: nonstructural protein NS4a *status predicted <N4A>
F:1863-2013/Product: nonstructural protein NS4 *status predicted <N8A>
F:2014-3010/Product: nonstructural protein NS5 *status predicted <NS5
F:2014-3010/Product: nonstructural protein S55,556,576,623,645/Binding site: carbohydrate
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          A; Molecule type: genomic RNA
A; Residues: 1-547, T', 549-621, V', 623-624, 'S', 626-652, 'DL', 655-761, 'T', 763-782
A; Cross-references: EMBL:X61591
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ATQTFLATSINGVLWTVYHGAGTRTIASPKGPVTQMYTNVDKDLVGWQAPQGSRSLTPCT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1065 ATQSFLATCVNGVCWTVYHGAGSKTLAGPKGPINQMYTNVDQDLVGWQAPSGAASLTPCT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CGSSDLYLVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              3 KKGSVVIVGRIN-----LSGDTAYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVST
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          DB 1; Length 3010;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               genome polyprotein - hepatitis C virus (isolate EUH1480)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              79.9%; Score 812.5; DB 1;
76.0%; Pred. No. 1.9e-65;
Live 20; Mismatches 20;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             174 STRGVAKAVDFIPVESLETTMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ilarity 76.0%;
Conservative 2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Best Local Similarity
Matches 155; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1005
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        đ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      A.Molecule type: genomic RNA
A.Residues: 2650-2707 <AR2>
C.Superfamily: hepatitis C virus genome polyprotein
C.Superfamily: hepatitis C virus genome polyprotein
C.Superfamily: hepatitis C virus genome polyprotein
F.2.115/Product: capsid protein C *status predicted <ARE>
F.2.115/Product: major envelope protein E *status predicted <ARE>
F.300-723/Product: major envelope protein NS2 *status predicted <ARE>
F.300-723/Product: nonstructural protein NS2 *status predicted <ARS>
F.1007-1615/Product: hepacivirin *status predicted <ARS>
F.1312-1317/Region: uncleotide-binding motif R (P-loop)
                                                                                                                                                                                                                                 genome polyprotein - hepatitis C virus (strain J)
N; Contains: capsid protein C; envelope protein M; major envelope protein E; nonstructura
protein NS4a; nonstructural protein NS4b; nonstructural protein NS5
C; Species: hepatitis C virus
C; Species: hepatitis C virus
C; Species: hepatitis C virus
C; Accession: A39253; PS6086
R; Kato, N; Hijikata, M.; Ootsuyama, Y.; Nakagawa, M.; Ohkoshi, S.; Sugimura, T.; Shimot
R; Kato, N; Hijikata, M.; Ootsuyama, Y.; Nakagawa, M.; Ohkoshi, S.; Sugimura, T.; Shimot
A; Title: Molecular cloning of the human hepatitis C virus genome from Japanese patients
A; Reference number: A39253; MuID:91088550; PMID:2175903
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        genome polyprotein - hepatitis C virus (isolate JK1)
N;Contains: capsid protein C; envelope protein M; hepacivirin (EC 3.4.21.98) (nonstructu protein NS4a; nonstructural protein NS4b; nonstructural protein NS5 C;Species; hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             R;Kato, N.; Ohkoshi, S.; Shimotohno, K.
Proc. Jpn, Acad. 65B, 219-223, 1989
A;Title: Japanese isolates of the non-A, non-B hepatitis viral genome show sequence variance number: PS0085
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   F;1616-1862/Product: nonstructural protein NS4a #status predicted <N4A>
F;1863-2013/Product: nonstructural protein NS4b #status predicted <N4B>
F;2014-3010/Product: nonstructural protein NS5 #status predicted <NS5>
F;196,209,234,250,305,325,417,423,430,448,532,556,576,623,645,1213,1255,2041,2077,2240,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       114 CGSSDLYLVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           3 KKGSVVIVGRINLSGD-----TAYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVST
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1005 RRGKEILLGPADSFGEQGWRLLAPITAYSQQTRGLLGCIITSLTGRDKNQVDGEVQVLST
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ATQTFLATSINGVLWTVYHGAGTRTIASPKGPVTQMYTNVDKDLVGWQAPQGSRSLTPCT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           A;Molecule type: genomic RNA
A;Residues: 1-3010 <KAI>
A;Cross_references: GB:D90208; NID:g221610; PIDN:BAA14233.1; PID:g221611
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Length 3010;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    81.0%; Score 823.5; DB 1; 75.5%; Pred. No. 1.9e-66; tive 23; Mismatches 18;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CTRGVAKAVDFIPVESMETTMRSP 1208
                                    174 STRGVAKAVDFIPVESLETIMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Query Match
Best Local Similarity 75.5%;
Matches 154; Conservative 2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              F;1316-1319/Region: DEXH motif
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  æ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  RESULT 8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           염
ô
                                                                     셤
```

```
C;Species: hepatitis C virus
C;Date: 19-Way-2000 #sequence_revision 19-May-2000 #text_change 17-Nov-2000
C;Date: 19-Way-2000 #sequence_revision 19-May-2000 #text_change 17-Nov-2000
C;Accession: J01303
B;Okamoto, H.; Okada, S.; Sugiyama, Y.; Kurai, K.; Iizuka, H.; Machida, A.; Miyakawa J. Gen. Virol. 72, 2697-2704, 1991
A;Title: Nucleotide sequence of the genomic RNA of hepatitis C virus isolated from a A;Reference number: J01303; MUID:92044440; PMID:1658196
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   A.Molecule type: genomic RNA
A.Residues: 1-3033 < OKA>
A.Residues: 1-3033 < OKA>
A.Residues: 1-3033 < OKA>
A.Residues: 1-3033 < OKA>
A.Experimental source: isolate HC-06 from a Japanese individual
A.Experimental source: solate HC-06 from a Japanese individual
C.Superfamily: hepatilis C virus genome polyprotein
C.Superfamily: hepatilis C virus genome polyprotein: serine proteinse; tr. pr. 2-115/Product: capsid protein: hydrolase; p-10op; polyprotein: serine proteinse; tr. pr. 2-115/Product: capsid protein M *status predicted < NED>
F:16-191/Product: anyelope protein E *status predicted < NED>
F:34-1010/Product: nonstructural protein NSI *status predicted < NSI>
F:31-11619/Product: nonstructural protein NSI *status predicted < NSI>
F:1116-1321/Region: uncleotide-binding motif B
F:1320-1323/Region: DEXH motif
                                                                 A)Cross-references: GB:D10562; GB:D90518; NID:g221523; PIDN:BAA01418.1; PID:g221524
G:Superfamily: hepatitis C virus genome polyprotein
G:Keywords: ATP: capsid protein: genome polyprotein
G:Keywords: ATP: capsid protein: genome polyprotein; hydrolase; nonstruc
G:Keywords: ATP: capsid protein C #status predicted <CPC>
F:1-115/Product: capsid protein C #status predicted <CPC>
F:1-115/Product: major envelope protein B #status predicted <NED>
F:192-189/Product: major envelope protein R*status predicted <NED>
F:30-733/Product: nonstructural protein NSI *status predicted <NSI>F:1011-1619/Product: nonstructural protein NSI *status predicted <NSI>F:1214-1241/Region: nucleotide-binding motif A (P-loop)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            F:130-133/Region: DEXH motif
F:1620-1866/Product: nonstructural protein NS4a #status predicted <N4A>
F:1867-2017/Product: nonstructural protein NS4b #status predicted <N4B>
F:2018-3013/Product: nonstructural protein NS5 #sratus predicted <NS5>
F:2018-3013/Product: nonstructural protein NS5 #sratus predicted <NS5>
F:196,209,233,299,305,417,423,430,448,477,534,542,558,578,627,649,1091,1217,1259,203
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         genome polyprotein - hepatitis C virus (isolate HC-J6)
N:Contains: capsid protein C; envelope protein M; hepacivirin (EC 3.4.21.98) (nonstr
protein NS4a; nonstructural protein NS4b; nonstructural protein NS5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      F;1620-1866/Product: nonstructural protein NS4a #status predicted <N4A>
F;1867-2017/Product: nonstructural protein NS4 #status predicted <NAB>
F;2018-3033/Product: nonstructural protein NS5 #status predicted <NOS-
F;2018-3033/Product: nonstructural protein NS5 #status predicted <NOS-
F;196,209,234,305,325,417,423,430,448,477,534,542,558,578,627,649,1091,1217,1259,203
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1153
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     79 IASPKGPVTQMYTNVDKDLVGWQAPQGSRSLTPCTCGSSDLYLVTRHADVIPVRRRGDSR 138
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         78
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       19 TAYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVSTATQTFLATSINGVLWTVYHGAGTRT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Score 674; DB 1; Length 3033;
Pred. No. 8.4e-53;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   66.4%; Score 675; DB 1; Length 3033;
69.8%; Pred. No. 6.8e-53;
tive 24; Mismatches 30; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           27; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        66.3%; Score 674;
68.7%; Pred. No. 8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Query Match
Best Local Similarity 69.8%
Matches 125; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Best Local Similarity 68.79
Matches 123; Conservative
                                                Š.
A; Molecule type: mRNA
A; Residues: 2678-2729
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   RESULT 11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               QQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       genome polyprotein - hepatitis C virus (strain HC-JB)
N:Contains: capsid protein C: envelope protein M: hepacivirin (EC 3.4.21.9B) (nonstructur protein NS4a; nonstructural protein NS4b; nonstructural protein NS5 (Species: hepatitis C virus (C:Species: hepatitis C virus (C:Species: hepatitis C virus (C:Accession: A40250; po0397; po0559
R:Okamoto, H.; Kurai, K.; Okada, S.I.; Yamamoto, K.; Lizuka, H.; Tanaka, T.; Fukuda, S.; Airology 18B, 331.341, 1992
A:Title: Full-length sequence of a hepatitis C virus genome having poor homology to repc A:Reference number: A40250; MUID:92230232; PMID:1314459
                                                                                                         A: Experimental source: genotype 5a, which predominates in South Africa
A:Note: the translation of the nucleotide sequence is not complete in this paper
A:Note: the translation of the nucleotide sequence is not complete in this paper
C: Superimanly: hepatitis C virus genome polyprotein
C: Superimanly: hepatitis C virus genome polyprotein
C: Reywords: ATP: glycoprotein: hydrolase; nucleotide binding; P-loop; polyprotein: p: 2-115/Product: capsid protein C *status predicted <CPC>
F: 192-389/Product: envelope protein E *status predicted <WED>
F: 384-408/Region: hypervariable *status predicted constructural protein NSI *status predicted <NSI>
F: 300-730/Product: nonstructural protein NSI *status predicted <NSI>
F: 1008-1616/Product: nonstructural protein NSI *status predicted <NSI>
F: 1018-1318/Region: nucleotide-binding motif B
F: 1313-1318/Region: nucleotide-binding motif B
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   A; Experimental source: isolate E-bl2
R; Kato, N.; Octsuyama, Y.; Ohkoshi, S.; Nakazawa, T.; Mori, S.; Hijikata, M.; Shimotohnq
B; Kato, N.; Octsuyama, Y.; Ohkoshi, S.; Nakazawa, T.; Mori, S.; Hijikata, M.; Shimotohnq
B; Aschem. Blophys. Res. Commun. 181, 279-285, 1991
A; Title: Distribution of plural HCV types in Japan.
A; Reference number: PQ0554; MUID: 92068204; PMID: 1720309
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                A; Molecule type: genomic RNA
A; Residues: 1-3033 < COKA>
A; Residues: 1-3033 < COKA>
A; Residues: 1-3033 < COKA>
A; Cross-references: GB:D10988; GB:D01221; NID:9221608; PIDN:BAA01761.1; PID:9221609
B; Chan, S.W.; McCmish, F.; Holmes, E.C.; Dow, B.; Peutherer, J.F.; Follett, E.; Yap, P. J. Gen, Virol, 73, 1131-1141, 1992
A; Title: Analysis of a new hepatitis C virus type and its phylogenetic relationship to A; Reference number: PQ0393; MUID:92268871; PMID:1316939
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1006 RRGREIFLGPADDIKTSGWRLLAPITAYAQQTRGVLGAIVLSLTGRDKNEAEGEVQFLST 1065
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1126 CGSADLYLVTRHADVIPARRRGDTRASLLSPRPISYLKGSSGGPIMCPSGHVVGVFRAAV 1185
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       114 CGSSDLYLYTRHADVIPVRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAV 173
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           23
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       3 KKGSVVIVGRIN-----LSGDTAYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVST
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         54 AIQIFLAISINGVLWTVYHGAGTRIIASPKGPVTQMYINVDKDLVGWQAPQGSRSLIPCI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              F.1677-1863/Product: nonstructural protein NS4a #status predicted <N4A>F.1864-2014/Product: nonstructural protein NS4b #status predicted <N4B>F.2015-3014/Product: nonstructural protein NS5 #status predicted <NS5>F.2010-2249/Region: interferon sensitivity determining #status predicted
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Length 3014;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            30; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                73.1%; Score 743.5; DB 1; 68.6%; Pred. No. 3.8e-59;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   25; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1186 CTRGVAKALEFVPVENLETTMRSP 1209
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             STRGVAKAVDFIPVESLETTMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             A;Molecule type: genomic RNA
A;Residues: 2678-2754 <CHA>
A;Cross references: DDBJ:D10134
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Ouery Match
Best Local Similarity 68.6%
Matches 140; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    F;1317-1320/Region: DEXH motif
                                   A;Residues: 1-3014 <CHA>
A;Cross-references: GB:Y13184
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          A; Accession: PQ0397
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               RESULT 10
```

ò 合 à å ò 윱

g Š

Qy         53 TATOTFLATSINGVLWTVYHGAGTRTIASPKGPVTQMYTNVDKDLVGWQAPQGSRSLTPC 112           Db	RESULT 14 B81104 nitrate/nitrite sensor protein (EC 2.7.3) NMB1249 [similarity] - Neisseria meningit C;Species: Neisseria meningitidis C;Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 19-Jan-2001 C;Accession: B81104 C;Accession: B81104 C;Accession: B1104 C;Accession: B104 C;Accession: B1104 C;Access	erences tal sou 1249 ly: nit autople	Query Match 9.1%; Score 92.5; DB 2; Length 590; Best Local Similarity 21.3%; Pred. No. 1.5; Matches 46; Conservative 26; Mismatches 79; Indels 65; Gaps 6; Qy 28 EEGCQETSQTGRDKNQVECKYQIVSTATQTFLATSINGVLWTVYHGAGTRIASFKGPVT 87 Db 213 EGGTPERQVGRCFNQMGGRLXILYDDLEGQVARQTRSLEKONQNLT 259	TPCTCGSSDL         : RVCLDGGSDV GPLLCPAGHA       RILLQTLGRQ 192	RESULT 15 C81911 nitrate/nitrite sensor protein (EC 2.7.3) NWA1418 [similarity] - Neisseria meningit C;Species: Neisseria meningitidis C;Species: O5-May-2000 *sequence_revision 05-May-2000 *text_change 02-Feb-2001 C;Accession: C81911 R;Parkhill, J.; Achtman, M.; James, K.D.; Bentiey, S.D.; Churcher, C.; Klee, S.R.; Mo ; HOLTOY, S.; Jagels, K.; Leather, S.; Moule, S.; Mungall, K.; Quall, M.A.; Rajandre Nature 404, 502-506, 2000 A;Title: Complete DNA sequence of a serogroup A strain of Neisseria menigitidis 22491 A;Reference number: A81775; MUID:20222556; PMID:10761919 A;Status: preliminary
Qy         19 TAYAOOTRGEEGCQETSQTGRDKNQVEGEVQIVSTATQTFLATSINGVLWTVYHGAGTRT 78           Db         1034 TAYAQOTRGLGTIVVSMTGRDKTEQAGEIQVLSTVTQSFLGTTISGVLWTVYHGAGNKT 1093           Qy         79 IASPKGPVTQNYTNVDKDLVGWQAPQGSRSLTPCTGGSSDLYLVTRHADVIPVRRGDSR 138           1	RESULT 12 T08841 polyprotein - douroucouli hepatitis GB virus A CiSpeciaes: douroucouli hepatitis GB virus A CiSpeciaes: douroucouli hepatitis GB virus A CiSpeciaes: douroucouli hepatitis GB virus A CiDate: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 17-Nov-2000 C;Accession: T08841 S:Erker. JC:; Desai, S.M.; Leary, T.P.; Chalmers, M.L.; Montes, C.C.; Mushahwar, I.K. J. Gen. virol. 79, 41-45, 1998 A;Itle: Genomic analysis of two GB virus A variants isolated from captive monkeys. A; Reference number: 216486; MUID:98120818; PMID:9460920 A;Accession: T08841 A;Molecule type: mRNA A;Molecule type: mRNA A;Residues: 1-3005 CERRA A;Residues: 1-3005 CERRA A;Cross-references: EMBL:AF022425; NID:92828599; FIDN:AAC40502.1; FID:92828600 C;Superfamily: hepatitis C virus genome polyprotein C;Keywords: polyprotein	Ouery Match 24.5%; Score 249; DB 2; Length 3005; Best Local Similarity 33.5%; Pred. No. 5.4e-14; Matches 55; Conservative 30; Mismatches 69; Indels 10; Gaps 3; Qy 33 ETSQTGRDKNOVEGEVQIVSTATQTFLATSINGVLMTVYHGAGTRIIASPKGPVTQMYTN 92 1	OY 93 VDKDLVGWQAPQGSRSITPCTGGSSDLXLVTRHADVIPVRRRGDSRGSLLSPRDISYLKG 152  1055 PSDDVAVYPLPSGASCLEPCKGGTOSVWCIRNDGALCHGRLSKLVELDLFTEISDFRG 1112  OY 153 SSGGPLLCPAGHAVGIFRAAVSTRGVAKAVDFIPVES 189  113 SSGSPLLCPEGHAVGMA-VSYLHRGVKVTGVRXVFPWFLPKDS 1155	T08839 T08839 T08839 T08839 C;Species: marmoset hepatitis GB virus A C;Species: marmoset hepatitis GB virus A C;Date: 23-Jul-1999 #sequence_revision 23-Jul-1999 #text_change 17-Nov-2000 C;Date: 23-Jul-1999 #sequence_revision 23-Jul-1999 #text_change 17-Nov-2000 C;CAccession: T08839 R;Erker, J.C.; Desai, S.M.; Leary, T.P.; Chalmers, M.L.; Montes, C.C.; Mushahwar, I.K. J. Gen. Virol. 79, 41-45, 1998 A;Title: Genomic analysis of two GB virus A variants isolated from captive monkeys. A;Reference number: 216486; MUID:98120818; PMID:9460920	A;Status: translated from GB/EMBL/DDBJ A;Status: translated from GB/EMBL/DDBJ A;Molecule type: genomic RNA A;Residues: 1.2970 < CERK> A;Cross-references: EMBL:AF023424; NID:g2828597; PIDN:AAC40501.1; PID:g2828598 C;Superfamily: hepatitis C virus genome polyprotein C;Superfamily: hepatitis C;Superfamily: h

```
A; Molecule type: DNA
A; Residues: 1-590 < PAR>
A; Residues: 1-590 < PAR>
A; Residues: 1-590 < PAR>
A; Cross-references: GB: AL162755; GB: AL157959; NID: g7379742; PIDN: CAB84658.1; PID: g738007
A; Experimental source: serogroup A, strain Z2491
C; Genetics: MAIA148
C; Genetics: Specification C; Superfamily: nitrate/nitrite sensor protein narx
C; Reywords: autophosphorylation; phosphohistidine intermediate) #status predicted
```

Query Match	atch	Query Match 9.1%; Score 92.5; DB 2; Length 590;	9.18;	Score	92.5;	DB 2	.: .:	ength 59	0;		
Matches	4.	best both Similarity 21.3%, Figu. NO. 1.3, Matches 46; Conservative 26; Mismatches 79; Indels 65; Gaps	<b>tive</b> 2	6; M1	smatche	, ,	.6	Indels	: 59	Gaps	6;
٥y	28	28 EEGCQETSQTGRDKNQVEGEVQIVSTATQTFLATSINGVLWTVYHGAGTRIIASPKGPVT 87	RDKNOVEG	EVQIVS	TATOTFI	ATSIN	HCVLW	TVYHGAGT	RTIASI	PKGPVT	87
Db	213	213 EGGTPEFKQVGRCFNQMGGRLKILYDDLEGQVAEQTRSLEKQNQNLT 259	RCFNQMGG	RLKILY	PDELEGOV	'AEQ		L	H:: RSLEK(	TRSLEKQNONLT	259
٥y	88	88 QMYTNVDKDLVGWQAPQ127	GWOAPO		9	SRSLI	PCTC	GSRSLTPCTCGSSDLYLVTRHAD	TRHAD		127
qq	260	260 LLY-QTTRDLHQSYIPQQAAEHFLNRILPAVGADSGRVCLDGGSDVYVSIHHADCGTAAS 318	QSYIPQQA	AEHFLN	RILPAVG	ADSGF	«VCLD	GGSDVYVS	IHHAD	GTAAS	318
٥y	128	128VIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVSTR 176	VRRRGDSR	GSLLSP	RPISYLE	gssec	PLIC	PAGHAVGI	FRAAVS	STR	176
qq	319	319 DLGKYHEEIFPIEYQNETLGRLLLSFPNGISLDEDDRILLLGTLGRQLGVSLAGAKQEEEK 378		GRLLLS	FPNGISI	DEDDE	ZITEO		SLAGAI	ÇEEEK	378
QY	177	177	GVAKAVDFIPVESLET 192	VAKAVD	FIPVE	SLET	192				
qq	379	379 RLLAVLQERNLIAQGLHDSIAQALTFLNLQVQMLET 414	IAQGLHDS	IAQALT	SIAQALTFLNLQVQMLET	MLET	414				

Search completed: August 30, 2003, 19:20:30 Job time: 17.2134 secs

```
GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.
```

OM protein - protein search, using sw model

August 30, 2003, 18:01:52; Search time 9.75674 Seconds (without alignments) 949.524 Million cell updates/sec Run on:

US-09-965-594-18 1017 1 MKKKGSVVIVGRINLSGDTA.....VAKAVDFIPVESLETIMRSP 197 Title: Perfect score: Sequence:

BLOSUM62 Gapop 10.0 , Gapext 0.5 Scoring table:

127863 seqs, 47026705 residues Searched:

127863 Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0% Maximum Match 100% Listing first 45 summaries

SwissProt\_41:\* Database :

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

88.84.5 88.84.5 88.84.5 88.34.5 87.55 87.55 87.55 77.7 77.			*			SUMMARIES	
1         854.5         84.0         3011         POLG_HCVH         P26664         h genome           2         848.5         83.4         3011         POLG_HCVTW         P2958         h genome           4         827.5         81.4         3010         POLG_HCVTW         P29646         h genome           5         823.5         81.0         3010         POLG_HCVJA         P26662         h genome           6         823.5         81.0         3010         POLG_HCVJA         P26663         h genome           7         66.3         3033         1 POLG_HCVJA         P26661         h genome           8         81.0         3010         1 POLG_HCVJA         P26661         h genome           9         87         8.6         3033         1 POLG_HCVJA         P26661         h genome           10         85.5         8.4         209         1 POLG_HCVJA         P26661         h genome           11         85.5         8.4         209         PADAD_PSEAR         Q98e110         P26661         h genome           12         8.1         1 POLG_TREVM         P26661         h genome         P26661         h genome           13         1 POL	Result	Score	Query Match		DB		Description
848.5         83.4         3011         POLG_HCVTW         POLG_HCTW         POLG_HCVTW         <	7	854.5	:	3011	7	POLG_HCV1	h genome
8 87.5         82.4         3010         1 POLG_HCVJT         POLG_GHCWIT           1 827.5         81.4         3010         1 POLG_HCVJT         G00269         h genome           823.5         81.0         3010         1 POLG_HCVJT         POLG_GHCWB         P26662         h genome           675         66.4         3010         1 POLG_HCVJR         P26661         h genome           674         66.3         3033         1 POLG_HCVJR         P26661         h genome           9         87.5         81.4         1 POLG_HCVJR         P26661         h genome           9         87.5         82.1         1 HNDA_ARATH         P26661         h genome           9         87.5         82.2         437         1 DEGI_ARATH         P26660         h genome           1         83.5         8.2         437         1 DEGI_ARATH         P26661         p genome           2         83.5         8.2         452         1 AAPP_HUWAN         P266987         J genome           3         81.2         452         1 AAPP_HUWAN         P26686         h genome           4         7.5         7.6         1 TEFE_HORSE         P26666         h genome		848.5	83.	3011	-	POLG_HCVH	h genome
8 22.5         81.4         3010         I POLG_HCVJT         Q00269         h genome PoLG_HCVJR         POLG_HCVJR         POLG_HCVJR         POLG_GB         h genome PoLG_HCVJR         POLG_HCVJR         POLG_GB         h genome PoLG_HCVJR         POLG_HCVJR         POLG_GB         h genome PoLG_HCVJR         POLG_HCVJR         POLG_HCVJR         POLG_GB         h genome PoLG_HCVJR         POLG_HCVJR         POLG_HCVJR         POLG_GB         h genome PoLG_GB         POLG_HCVJR         POLG_HCVJR         POLG_HCVJR         POLG_HCVJR         POLG_HCVJR         POLG_GB         h genome PoLG_GB         POLG_HCVJR         POLG_HCVJR         POLG_GB         POLG_GB         h genome PoLG_GB		837.5	82	3010	-	POLG_HCVTW	h genome
8 83.5         81.0         3010         I POLG_HCVJA         P26663         H genome           673         66.4         3031         I POLG_HCVJA         P26663         H genome           674         66.3         3033         I POLG_HCVJA         P26661         H genome           87.5         66.4         3033         I POLG_HCVJA         P26661         H genome           9         87.5         8.4         209         I HHOA_MRATH         Q98e17         arabidop           9         85.5         8.4         209         I DEGI_ARATH         Q98e17         arabidop           1         83.5         8.2         437         I DEGI_ARATH         Q95e17         arabidop           2         83.4         4         209         I DEGI_ARATH         Q96e16         pacudomo           3         8.2         4.37         I DEGI_ARATH         Q13685         homosap           4         7.9         3414         I POLG_TEBEVM         P21436         t genome           5         7.7         76         I TEBL_NEMB         P0156         nesseria           7         7.6         1.7         341         I POLG_TEBEV         P07520         t genome <td></td> <td>827.5</td> <td>81.</td> <td>3010</td> <td>7</td> <td>POLG_HCVJT</td> <td>h genome</td>		827.5	81.	3010	7	POLG_HCVJT	h genome
6 33.5         81.0         3010         POLG_HCVJA         P26662 h genome           675         66.4         3033         1 POLG_HCVJB         P26661 h genome           9 87         66.4         3033         1 POLG_HCVJB         P26661 h genome           9 87.5         8.4         66.3         3034         1 HOA_ARATH         P26661 h genome           1 83.5         8.2         437         1 DEGI_ARATH         022609 arabidop           2 83.6         8.2         437         1 DEGI_ARATH         022609 arabidop           3 83.6         8.2         437         1 DEGI_ARATH         022609 arabidop           3 83.7         8.2         452         1 AAMP_HUMAN         013685 home sap           3 7.8         7.7         7.6         1 TREE_HORSE         P27425 equus           4 8.5         7.7         764         1 TREE_HORSE         P27425 equus           5 7.7         764         1 TREE_HORSE         P01600000         P0160000           6 7.7         312         1 POLG_TREVS         P01720         p010000           7 7         76         1 TREL_BDVSE         P016180         P016200           7 7         7.6         341         1 NTR_SPTOL         P016314 <td></td> <td>823.5</td> <td>81.</td> <td>3010</td> <td>-</td> <td>POLG_HCVBK</td> <td>h genome</td>		823.5	81.	3010	-	POLG_HCVBK	h genome
675 66.4 3033   POLG_HCVJ6   P26661 h genome		823.5	81.	3010	~	POLG_HCVJA	h genome
674 66.3 3033   POLG HCVJ6   P26660 h genome   685.8 8.4 209   PAAD_PSEAE   Q99e17 arabidop   685.8 8.4 209   PAAD_PSEAE   Q99e17 arabidop   G99e17 arabidop   G99e18   G99e17 arabidop   G99e18   G99e		675	.99	3033	н	POLG_HCVJ8	h genome
9 87. 8.6 321 1 HHOA.ARATH Q9sel7 arabidop 0 1 PAAD_PEBAR G95.6 9 seld arabidop 0 226.09 arabidop 0 22		674	99	3033	٦	POLG_HCVJ6	h genome
85.5   8.4   209   PAAD_PSEAE   Q9hx0B pseudomo   83.5   8.2   452   1 DeG_ARATH   Q136B5 homo sapone   Q2609 arabidopo   Q2609 arabidop	6	87	80	321	-	HHOA ARATH	
1 83.5 8.2 437 1 DEGI_ARATH 022609 arabidop 22 83 8.2 452 1 AAMP_HUWAN 013685 homo sap 23 79.5 7.8 3414 1 POLG_TBEVW P29837 1 genome 24 78.5 7.7 766 1 TREE_HORSE 008180 drosophila or 26 7.7 7 76 1 TREE_HORSE 008180 drosophila or 27 7 7 7 6 1 TREE_HORSE 008180 drosophila or 27 7 7 7 6 1 TREE_HORSE 008180 drosophila or 28 7 7 7 7 6 1 TREE_HORSE 008180 drosophila or 28 7 7 7 7 6 1705 1 POLG_TBEVS 00550 meisseri 0077 7 7 6 1705 1 POLG_TBEVS 00550 meisseri 0077 7 7 6 1705 1 POLG_TBEVS 00550 meisseri 0077 7 7 6 1705 1 POLG_TBEVS 005314 spinacia 00765 7 7 7 6 1705 1 POLG_TBEV 001299 drosophila or 27 7 7 6 1705 1 POLG_TBEV 001299 drosophila or 28 7 7 7 7 6 1705 1 POLG_TBEV 001299 drosophila or 28 7 7 7 7 8 1 1 POLG_TBEV 001299 drosophila or 28 7 7 7 7 8 1 1 POLG_TBEV 001314 y genome 00765 7 7 7 7 8 1 1 POLG_TEV 001201 Drosophila or 28 7 7 7 7 8 1 1 POLG_TEV 001201 Drosophila or 28 7 7 7 7 8 1 1 POLG_TEV 001201 Drosophila or 28 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7	10	85.5	ω,	209	-	PAAD_PSEAE	
2         83         8.2         452         J AAMP_HUWAN         Q13685 homo sap           3         79.5         7.8         3414         1 POLG_TMBVW         P29837 J genome           5         7.8         3414         1 POLG_TMBVW         P14336 t genome           6         7.8         34.1         1 TEE_HORE         P2425 equus ca           6         7.7         764         1 CCR_DMME         Q08180 drosophia           7         7.6         1 TB1_NEIMB         Q08056 neisseria           8         7.7         3412         1 POLG_TBEVS         P07720 t genome           9         7.7         5 6.5         1 NIR_SPPOL         P07314 spinam ansc           9         7.7         6 594         1 NIR_SPPOL         P05208 mus musc           1         7.6         534         1 POLG_TEEVH         P01299 mus musc           2         7.6         341         1 POLG_TEEVH         P01299 mus musc           3         76.5         7.5         333         MCA_RHIME         P013931 secheric           6         7.5         341         1 POLG_TEEVI         P013314 y genome           7         7.5         341         1 POLG_TEEVI         P013314 y genome	11	83.5	в.	437	-	DEG1_ARATH	
3         79.5         7.8         3414         1 POLG_LANVT         PE29837 1 genome           4         79         7.8         3414         1 POLG_TERV         POLG_TERV         PO14336 t genome           5         7.8         7.7         764         1 ITRE_LORSE         POLG_TERV         POLG TERV         Q09180 drosophi           7         76         7.7         3412         POLG_TERVS         POT720 t genome         POT720 t genome           9         7.7.5         7.6         263         I GRAK_MOUSE         POT720 t genome         POT3280 mus musc           0         7.7         7.6         594         1 NIR_SPTOL         POT3314 spinacia           1         7.7         7.6         3414         POLG_TERVH         PO1299 t genome           1         7.7         7.6         341         POLG_TERVH         PO1497 squirrel           3         7.6.5         7.5         333         1 WOSA_RHIME         PO1407 squirrel           4         7.6.5         7.5         3411         POLG_TERV         PO1407 squirrel           5         7.5         3411         POLG_TERV         POLG_TERV           6         7.5         3411         POLG_TERV         POLG_TER	15	83	æ	452	٦	AAMP_HUMAN	
4         79         7.8         3414         1 POLG_TBEVW         P14336 t genome           5         7.7         76         1 TRE_HORE         P2425 equus cand cand cand cand cand cand cand cand	13	79.5	7.	3414	7	POLG_LANVT	l genome
5         78.5         7.7         706         1         TRFE_HORSE         P27425 equus ca           6         78.5         7.7         764         1         TCCR_DROME         008180 drosophia           7         7.6         7.7         3412         1         POLG_TBEVS         P07720 t genome           9         7.7         5         6         594         1         NTR_SPPOL         P053205 mus	7	79	7	3414	-	POLG_TBEVW	t genome
6 78.5 7.7 764 1 ICCR_DROME 008180 drosophi   78 7.7 3412 1 PDLG_TBENB 009056 neisser1   78 7.7 3412 1 PDLG_TBENB 009056 neisser1   8 7.7 3412 1 PDLG_TBENB   9 77.5 7.6 263 1 GRAK_MOUSE	12	78.5	7.	706	-	TRFE_HORSE	
7 / 8         7.7         911         1 TB11_NEIMB         Q09056 neisseri           8 / 7.6         263         1 GRAK_MOUSE         P07720 t genome           9 / 7.5         7.6         594         1 NIR_SPIOL         P05314 spinacia           1 / 7         7.6         594         1 NIR_SPIOL         P05314 spinacia           1 / 7         7.6         3414         1 POLG_TBEVH         P01299 mus musc           2 / 7         7.6         3414         1 POLG_TBEVH         P01299 t genome           3 / 76.5         7.5         323         1 VPRT_SMRVH         P01407 squirrel           4 / 5.5         7.5         3431         1 POLG_TEEVH         P01407 squirrel           5 / 6.5         7.5         341         1 POLG_YEEVU         P03314 y genome           6 / 7.5         341         1 POLG_YEEVU         P03314 y genome           7 / 6         7.5         341         1 POLG_YEEVU         P03314 y genome           8 / 7.5         7.4         485         1 X136_TREPA         OB1317 treponem           9 / 7.5         7.4         2269         WX1B_BOVIN         O28142 bos taur           1 / 75         7.4         467         1 WX1B_BOVIN         O28142 bos taur	16	78.5	7.		Н	ICCR_DROME	
8         7.7         3412         1 POLG_TEBVS         P07720 t genome           9         77.5         7.6         594         1 NIR_SPLOL         P053205 mus musc           1         7.6         594         1 NIR_SPLOL         P053205 mus musc           2         7.7         7.6         1705         1 PPPO_MOUSE         P05314 splanes musc           2         7.7         7.6         1705         1 VPRT_SWRVH         P01299 t genome           4         76.5         7.5         333         1 WORA_RIME         P01607 raizobiu           5         7.5         341         1 POLG_TEEV         P03314 y genome           6         7.5         341         1 POLG_TEFV         P03314 y genome           7         7.5         341         1 POLG_TEFV         P03931 y genome           8         7.5         341         1 POLG_TEFV         P019901 y genome           8         7.5         7.4         485         1 VISC_TEFV         P019901 y genome           8         7.5         7.4         485         1 VISC_TEFV         P019901 y genome           9         7.5         7.4         467         1 NXIB_BUNIN         Q08112         D01001 y genome	17	78	7.			TB11_NEIMB	
9 77.5 7.6 263 1 GRAK_MOUSE 035205 mus musc 0 77 7.6 594 1 NIR_SPIOL P05314 spinacia 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2	18	78	7.		-	POLG_TBEVS	
0         77         7.6         594         1 NIR_SFIOL         PO5314 spinacia           1         7.7         7.6         3414         1 POLG_TBEVH         PO1299 mus musc           3         7.6.5         7.5         323         1 VPRT_SMRVH         PO14079 t qenome           4         76.5         7.5         33.3         1 MOSA_RHIME         PO1407 rhizobiu           5         76.5         7.5         45.2         1 MLTD_ECOLI         PO39314 y genome           6         7.5         3411         1 POLG_EEFV         PO1314 y genome           7         7.6         7.5         34.1         1 POLG_EFFV           8         7.5.         7.4         485         1 Y136_TREPA         O81372           9         7.5.         7.4         2269         1 WDR9_HUMAN         Q9nsi6 homo sapp           1         7.5         7.4         467         1 NX1B_BOVIN         Q9nsi6 homo sapp           1         7.5         7.4         467         1 WIB_HUMAN         Q9psi3 homo sapp           1         7.5         7.4         165         1 POL_GALV         P21414 glbbon a           1         7.5         7.4         165         1 GRAD_MOUSE	19	77.5	7.		-	GRAK_MOUSE	
1 77 7.6 1705 1 PTPO_MOUSE P70289 mus musc	20	77	7		-1	NIR_SPIOL	
2 77 7.6 3414 1 POLG_TBEVH 001299 t genome 76.5 7.5 323 1 VPRT_SMRVH 007607 squirrell MOSA_EMINE 007607 rhizobiu 76.5 7.5 334 1 MCTD_ECOLI P23931 escheric 6 7.5 3411 1 POLG_YEFV1 P19901 y genome 7 7.5 7.4 485 1 Y136_TREPA 093172 treponem 9 75.5 7.4 2269 1 WDR9_HUMAN 09025 homo sap 1 75 7.4 1165 1 POLG_ALV PROMINE 090253 homo sap 1 75 7.4 1165 1 POLG_ALV PROMINE 090253 homo sap 1 75 7.4 1165 1 POLG_ALV PROMINE 090253 homo sap 1 75 7.4 1165 1 POLG_ALV PROMINE 090253 homo sap 1 74.5 7.3 248 1 GRAD_MOUSE P11033 mus musc	21	77	7.		H	PTPO_MOUSE	
3         76.5         7.5         323         1 VPRT_SMRVH         P21407 squirrel           4         76.5         7.5         452         1 MLTD_ECOLI         P23931 escheric           5         76.5         7.5         3411 l POLG_YEFV1         P01314 y genome           7         76         7.5         3411 l POLG_YEFV2         P19901 y genome           8         75.5         7.4         485 l Y136_TREPA         081172 treponem           9         75.5         7.4         2269 l WDR9_HUMAN         091312 treponem           0         75         7.4         467 l WX1B_BOVIN         028142 bos taur           1         75         7.4         165 l POL_GALV         P21414 gibbon asp           1         75         7.4         165 l POL_GALV         P21414 gibbon asp           3         74.5         7.3         246 l GRAD_MOUSE         P11033 mus mus musc	22	77	7.		-	POLG_TBEVH	t genome
4         76.5         7.5         33.3         1 MOSA, RHIME         007607 Fhizobium           5         76.5         7.5         4.52         1 MLTD_ECOLI         P23931 escherich benation           7         7.5         34.11         1 POLG, YEFV2         P01991 y genome           8         75.5         7.4         485         1 Y136_TREPA         O83132 treponema           9         75.5         7.4         2269         1 WDR9, HUMAN         Q9nsi6 homo sapi           0         75         7.4         467         1 WIB_BOVIN         Q28142 bos tauru           1         75         7.4         467         1 POL_GALY         PDL_GALY           2         7.5         7.4         467         1 POL_GALY         PDL_GALY           3         7.4         97.3         1 POL_GALY         PDL_GALY           4         15         10.0         PDL_GALY         PDL_GALY <t< td=""><td>23</td><td>76.5</td><td>7</td><td></td><td>-</td><td>VPRT_SMRVH</td><td>squirrel</td></t<>	23	76.5	7		-	VPRT_SMRVH	squirrel
5         76.5         7.5         452         1 MLTD_ECOLI         P23331 escherich           6         7.5         3411         1 POLG_YEEV1         P03314 y genome           7         7.5         3411         1 POLG_YEEV2         P193314 y genome           9         75.5         7.4         485         1 X136_TREPA         083172 treponema           9         75.5         7.4         2269         1 WDR9_HUMAN         Q9nsi6 homo sapi           0         75         7.4         467         1 NX1B_BOVIN         Q28142 bos tauru           1         75         7.4         467         1 VP18_HUMAN         Q9p253 homo sapi           1         75         7.4         1165         1 POL_GALV         P21414 glabon app           2         75         7.4         1165         1 GRAD_MOUSE         P11033 mus muscu	24	76.5	7.		7	MOSA_RHIME	
6 76 7.5 3411 1 POLG_YEFV1 P03314 y genome 7 75.5 7.4 485 1 Y136_TREPA O83172 treponema 9 75.5 7.4 2269 1 WDR9_HUMAN O90816 homo sapi 0 75 7.4 467 1 WX1B_BOVIN O28144 box tauru 1 75 7.4 467 1 WX1B_BOVIN O90816 homo sapi 1 Y15 7.4 1165 1 POL_GALV POLGE	25	76.5	7.		Н	MLTD_ECOLI	-
7         76         7.5         3411         1         POLG_YEFV2         P19901 y genome           8         75.5         7.4         2485         1         Y136_TREPA         O83172 treponema           0         75.5         7.4         2269         1         WDR9_HUMAN         Q9nsi6 homo sapi           0         75         7.4         467         1         NX1B_BOVIN         Q28142 bos tauru           1         75         7.4         467         1         PPLB_GILMAN         P2133 homo sapi           1         75         7.4         1165         1         POL_GALV         P21414 glbbon app           3         74.5         7.3         248         1         GRAD_MOUSE         P11033 mus muscu	36	16	7.		٦	POLG_YEFV1	y genome
8         75.5         7.4         485         1         X136_TREPA         083172           9         75.5         7.4         4269         1         WDR9_HUMAN         Q9nsi6           0         75         7.4         467         1         NY1B_BOWIN         Q9R142           2         7.5         7.4         973         1         VP1B_HUMAN         Q9p253           2         75         7.4         1165         1         POL_GALV         P21414           3         74.5         7.3         248         1         GRAD_MOUSE         P11033	27	92	7.	,	~	POLG_YEFV2	1 y genome
9 75.5 7.4 2269 1 WTR9_HUMAN Q9nsi6 0 75 7.4 467 1 NATB_BOVIN Q28142 1 75 7.4 1913 1 VP18_HUMAN Q9p553 2 75 7.4 1165 1 POL_GALV P21414 3 74.5 7.3 248 1 GRAD_MOUSE P11033	28	75.5	7.		-	Y136_TREPA	~
0 75 7.4 467 1 NXIB_BOVIN 028142 1 75 7.4 973 1 VURL_HUMAN 099253 2 75 7.4 1165 1 POL_GALV P21414 3 74.5 7.3 248 1 GRAD_MOUSE P11033	53	75.5	۲.		7	WDR9_HUMAN	
1 75 7.4 973 1 VP18_HUMAN 09p253 2 75 7.4 1165 1 POL_GALV P21414 3 74.5 7.3 248 1 GRAD_MOUSE P11033	30	75	7.		-	NX1B_BOVIN	٠.
2 75 7.4 1165 1 POL_GALV P21414 91bb 3 74.5 7.3 248 1 GRAD_MOUSE P11033 mus	31	75	7	973	-	VP18_HUMAN	~
3 74.5 7.3 248 1 GRAD_MOUSE P11033 mus	32	75	7	7	-	POL_GALV	_
	33	74.5		248	7	GRAD_MOUSE	S mus

P26596 lactococcus P97608 rattus norv P54748 rattus norv P54856 ustilago ma Q96qs1 homo sapien Q8621 mus musculu P28651 oryctolagus Q28146 bos taurus Q28146 bos taurus Q28146 bos taurus	
MCP_BPF41 OPLA_RAT CNAA_RAT DPM1_GSTWA PHWX_HUWAN FXH1_MOUSE MWA1_BOVIN NX1A_BOVIN POLG POWVI	POS8_CAREL
<b>лелапалапа</b>	
301 1288 844 294 320 401 2499 2499 1530 3415	861
	7.1
44	72.5
488 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8	4.5

## ALIGNMENTS

```
PRT; 3011 AA
                                                                                                                                                                                                                                                       174 STRGVAKAVDFIPVESLETTMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SEQUENCE FROM N.A.
MEDLINE-92052256; PubMed-1658800;
                                                                                                                                                                                                                                                                                                                                                                            STANDARD:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Hepacivirus.
NCBI_TaxID=11108;
                                                                                                                                                                                                                                                                                                                                                                              POLG_HCVH
                                                                                                                                                                                                                                                                                                                                   RESULT 2
POLG_HCVH
                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                          셤
                                                                   ò
                                                                                                        8
                                                                                                                                                   à
                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CELLOLAR AMINOPEPTIDASE.
CAPSID PROTEIN C (POTENTIAL).
MATRIX PROTEIN (POTENTIAL).
MAJOR ENVELOPE PROTEIN E (POTENTIAL).
NONSTRUCTURAL PROTEIN NS1/R2 (POTENTIAL).
PROPEREAE/HELICASE NS3 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS48 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS48 (POTENTIAL).
RNA-DIRECTED RNA POLYMERASE (POTENTIAL).
RNA-DIRECTED RNA POLYMERASE (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Polyprotein; Glycoprotein; Transferase; RNA-directed RNA polymerase;
Core protein; Coat protein; Envelope protein; Helicase; ATP-binding;
Transmembrane; Nonstructural protein; Hydrolase; Serine protease;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SIMILARITY).
SIMILARITY).
SIMILARITY).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (POTENTIAL)
(POTENTIAL)
(POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  (POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ( POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          (POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Score 854.5; DB 1; Length 3011; Pred. No. 2.1e-71;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     REMOVED FROM CAPSID PROTEIN C
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CRC64:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               RELAY SYSTEM (BY
RELAY SYSTEM (BY
RELAY SYSTEM (BY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              GLCNAC, . .
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        65F8C9447FCE5AF9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   N-LINKED (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      N-LINKED (GLCNAC
N-LINKED (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               N-LINKED (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ATP (POTENTIAL)
DECH BOX.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       LINKED
                                                                                                   InterPro; IPR000745; HCV_NS4a.
InterPro; IPR001409; HCV_NS4b.
InterPro; IPR001409; HCV_NS4b.
InterPro; IPR001506; HCV_RdRP.
InterPro; IPR001509; Helicase_C.
InterPro; IPR001095; RNA_pol_DS_PS.
InterPro; IPR007094; RNA_pol_PSvir.
Pfam; PF01543; HCV_core; 1.
Pfam; PF01543; HCV_core; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CHARGE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CHARGE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CHARGE
                                                                                                                                                                                                                           InterPro; IPR007094; RNA_DOI_PSG
Pfan; PF01542; HCV_capsid; 1.
Pfan; PF01542; HCV_capsid; 1.
Pfan; PF01542; HCV_capsid; 1.
Pfan; PF01542; HCV_cnv; 1.
Pfan; PF01546; HCV_NS1; 1.
Pfan; PF01006; HCV_NS3; 1.
Pfan; PF01006; HCV_NS4; 1.
Pfan; PF01006; HCV_NS5a; 1.
Pfan; PF00201; HCV_NS5a; 1.
                                             HCV_NS1.
HCV_NS2.
HCV_NS3.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 84.0%;
82.8%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        AA; 327197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        3011
369
1083
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1165
1165
1237
1319
196
209
234
305
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            645
2041
2077
                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ProDom; PD186062; HCV_I
SMART; SM00487; DEXDC;
IPR002521;
                                                              IPR002518;
IPR004109;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     3D-structure.
INIT_MET
                                                              InterPro;
InterPro;
                          nterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CHAIN
TRANSMEM
ACT_SITE
ACT_SITE
ACT_SITE
NP_BIND
SITE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Ouery Match
Best Local S
Matches 169
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CHAIN
    Solution in the contract of th
```

3 KKGSVVIVGRIN-----LSGDTAYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVST

Gaps

6

Indels

9; Mismatches

Conservative

169;

ð

```
54 ATQTFLATSINGVLWTVYHGAGTRIIASPKGPVTQMYTNVDKDLVGWQAPQGSRSLTPCT 113
                                                                                                                                        114 CGSSDLYLVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAV 173
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   01-A0G-1992 (Rel. 23, Created)
01-A0G-1992 (Rel. 23, Last sequence update)
15-SEP-2003 (Rel. 42, Last annotation update)
16-SEP-2003 (Rel. 42, Last annotation update)
16-SOP-2003 (Rel. 42, Last annotation update)
16-SOP-2003 (Nall): Protein Pl (GP83) (GP85); Envelope glycoprotein E2
16-SOP-2003 (Nall): Protein Pl Nonstructural protein NS2 (P21)
16-SOP-3-1: Protease/hellcase NS3 (P70) (Hepacivirin)
16-SOP-3-1: Protein NSS (P86): Nonstructural protein
16-SOP-3-1: Nonstructural protein NSSA (P86): Nonstructural protein
16-SOP-3-1: Nonstructural Protein NSSA (P86): Nonstructural protein
16-SOP-3-1: Nonstructural Protein NSSA (P86): Nall NSSA (P86): Nonstructural Protein NSSA (P86): Nonstru
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                MEDLINE-98154321; PubMed-9493270; Kim J.L., Morgenstern K.A., Griffith J.P., Dwyer M.D., Thomson J.A., Murcko M.A., Lin C., Caron P.R.; "Hepatitis C virus NS3 RNA helicase domain with a bound oligonucleotide: the crystal structure provides insights into the mode of unwinding.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             H.V., Weber P.C.;
                                  Structure 6:89-100(1998).
-:- FUNCTION: PROTEASE NS2 IS RESPONSIBLE FOR THE CLEAVAGE OF NS2-NS3-1- FUNCTION: PROTEASE NS3 IS RESPONSIBLE FOR THE CLEAVAGE OF
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CATALYTIC ACTIVITY: Hydrolygis of four peptide bonds in the viral precursor polyprotein, commonly with Asp or Glu in the P6 position, Cys or Thr in P1 and Ser or Ala in P1.

CATALYTIC ACTIVITY: Nucleoside triphosphate +
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              NS3-NS4A, NS4A-NS4B, NS4B-NS5A AND NS5A-NS5B.
-!- FUNCTION: NS4A FORMS A COMPLEX WITH NS3 AND IS ESSENTIAL FOR THE ACTIVATION OF NS3.
-!- FUNCTION: NS5A SEEMS TO HAVE A TRANSCRIPTIONAL ACTIVATORY ROLE.
-!- FUNCTION: NS5B IS A RNA-DEPENDENT RNA POLYMERASE THAT PLAYS AN ESSENTIAL ROLE IN THE VIRUS REPLICATION.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS: E1 AND E2. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MRNA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 U
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Prince A.M.; "Genomic strain H of hepatitis "Genomic structure of the human prototype strain H of hepatitis virus: comparison with American and Japanese isolates."; Proc. Natl. Acad. Sci. U.S.A. 88:10292-10296(1991).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Inchauspe G., Zebedee S., Lee D.H.H., Sugitani M., Nasoff M.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       MEDLINE-97331322; PubMed-9187654;
Yao N., Hesson T., Cable M., Hong 2., Kwong A.D., Le H.V
"Structure of the hepatitis C virus RNA helicase domain.
Nat. Struct. Biol. 4:463-467(1997).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         X-RAY CRYSTALLOGRAPHY (2.1 ANGSTROMS) OF 1207-1657.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS) OF 1192-1657.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (RNA) (N).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ---
```

```
(POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Length 3011;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      327142 MW; 772CBB29CCD94753 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Score 848.5; DB 1;
Pred, No. 7.5e-71;
                                  (GLCNAC.
(GLCNAC.
(GLCNAC.
(GLCNAC.
(GLCNAC.
(GLCNAC.
              GLCNAC.
             N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          83.4%;
                                                                                                                                                                                                                                                                                                                                                                                                                                         11411
1417
1417
1436
11436
11453
11463
11463
11501
11501
11511
11511
11554
11564
11564
11564
                                                                                                                                        1385
1393
1409
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         580
597
598
611
618
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Ä
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Query Match
Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                      14410
14444
14431
14411
14556
14611
14611
14611
15111
15111
15111
15111
15111
15111
15111
15111
15111
15111
15111
15111
15111
15111
15111
15111
15111
15111
15111
15111
15111
15111
15111
15111
15111
15111
15111
15111
15111
15111
15111
15111
15111
15111
15111
15111
15111
15111
15111
15111
15111
15111
15111
15111
15111
15111
15111
15111
15111
15111
15111
15111
15111
15111
15111
15111
15111
15111
15111
15111
15111
15111
15111
15111
15111
15111
15111
15111
15111
15111
15111
15111
15111
15111
15111
15111
15111
15111
15111
15111
15111
15111
15111
15111
15111
15111
15111
15111
15111
15111
15111
15111
15111
15111
15111
15111
15111
15111
15111
15111
15111
15111
15111
15111
15111
15111
15111
15111
15111
15111
15111
15111
15111
15111
15111
15111
15111
15111
15111
15111
15111
15111
15111
15111
15111
15111
15111
15111
15111
15111
15111
15111
15111
15111
15111
15111
15111
15111
15111
15111
15111
15111
15111
15111
15111
15111
15111
15111
15111
15111
15111
15111
15111
15111
15111
15111
15111
15111
15111
15111
15111
15111
15111
15111
15111
15111
15111
15111
15111
15111
15111
15111
15111
15111
15111
15111
15111
15111
15111
15111
15111
15111
15111
15111
15111
15111
15111
15111
15111
15111
15111
15111
15111
15111
15111
15111
15111
15111
15111
15111
15111
15111
15111
15111
15111
15111
15111
15111
15111
15111
15111
15111
15111
15111
15111
15111
15111
15111
15111
15111
15111
15111
15111
15111
15111
15111
15111
15111
15111
15111
15111
15111
15111
15111
15111
15111
15111
15111
15111
15111
15111
15111
15111
15111
15111
15111
15111
1511
15111
15111
15111
15111
15111
15111
15111
15111
15111
15111
15111
15111
15111
15111
15111
15111
15111
15111
15111
15111
15111
15111
15111
15111
15111
15111
15111
15111
15111
15111
15111
1511
15111
15111
15111
15111
15111
15111
15111
15111
15111
15111
1511
15111
15111
15111
15111
15111
15111
15111
15111
15111
15111
15111
15111
15111
15111
15111
15111
15111
15111
15111
15111
15111
15111
15111
15111
15111
15111
15111
15111
15111
15111
15111
1511
15111
1511
1511
1511
15111
15111
15111
15111
15111
15111
15111
1
    1317
1323
1336
1343
1352
1352
1368
1368
1373
1376
                                                                                                                                                                                                                                                                                                                                                                                                         382
1389
1397
                                  CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     HELIX
HELIX
STRAND
HELIX
HELIX
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    STRAND
STRAND
STRAND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        TURN
HELIX
TURN
HELIX
TURN
STRAND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                STRAND
    This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the Burnepan Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                  SIMILARITY).
SIMILARITY).
SIMILARITY).
PTM: THE STRUCTURAL PROTEINS C, E1 AND E2 ARE PRODUCED BY PROTEOLYTIC PROCESSING BY THE HOST SIGNAL PEPTIDASES. SIMILARITY: THE NS2 PROTEASE BELONGS TO PEPTIDASE FAMILY U39. SIMILARITY: THE NS3 PROTEASE BELONGS TO PEPTIDASE FAMILY S29.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             LINKED (GLCNAC. .) (POTENTIAL).
LINKED (GLCNAC. .) (POTENTIAL).
LINKED (GLCNAC. .) (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    REMOVED FROM CAPSID PROTEIN C BY CELLULAR AMINOPEPTIDASE.
CAPSID PROTEIN C.
ENVELOPE GLYCOPROTEIN E1.
ENVELOPE GLYCOPROTEIN E2.
PROTEIN P7.
NONSTRUCTURAL PROTEIN NS2.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          PROTEASE/HELICASE NS.
NONSTRUCTURAL PROTEIN NS4A.
NONSTRUCTURAL PROTEIN NS4B.
NONSTRUCTURAL PROTEIN NS5B.
POTENTIAL.
CHARGE RELAY SYSTEM (BY SIMII
CHARGE RELAY SYSTEM (BY SIMII
ATP (POTENTIAL).
                                                                                                                                  EMBL, M67463; AAA4554.1; --
PIR, A36814; GWWVCH,
PDB; 1HEI; 25.NOV-98.
PDB; 1A1V; 16-FEB-99.
PDB; 1A1N; 17-JUN-98.
MEROPS; S29 (001): --
MEROPS; U39 (001): --
TRANSFAC; T04155; --
InterPro; IPR001241; DEAD.
InterPro; IPR002522; HCV_capsi
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          191
383
383
746
809
1026
11711
11711
11711
1107
1108
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    192
384
747
810
11027
11027
11107
11083
11107
11165
111165
1230
1316
234
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           3D-structure.
INIT_MET
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               TRANSMEM
ACT_SITE
ACT_SITE
ACT_SITE
NP_BIND
SIND
CARBOHYD
CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CHAIN
CHAIN
CHAIN
CHAIN
CHAIN
CHAIN
```

18; Indels

Matches 167; Conservative

```
3D-structure.
INIT_MET
                                      InterPro; InterP
                                                                                                                                                                     (aterPro;
                                                                                                                                                                                                           nterPro;
                                                                                                                                  InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ACT_SITE
ACT_SITE
NP_BIND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          TRANSMEM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ACT_SIT!
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CHAIN
CHAIN
CHAIN
CHAIN
CHAIN
                                                                                                                                                                                                                              Pfam;
     This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation-the Buropean Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/
                                                         54 ATQTFLATSINGVLWTVYHGAGTRTIASPKGPVTQMYTNVDKDLVGWQAPQGSRSLTPCT 113
                                                                                                                                                                                      114 CGSSDLYLVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAV 173
                                      3 KKGSVVIVGRIN-----LSGDTAYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVST 53
                                                                                                                                                                                                                                                                                                                                                                                                                     01-APR-1993 (Rel. 25, Created)
01-APR-1993 (Rel. 25, Last sequence update)
15-5P-2003 (Rel. 42, Last annotation update)
15-5P-2003 (Rel. 42, Last annotation update)
6enome polyprotein [Contains: Capsid protein C (Core protein) (P22);
Envelope glycoprotein El (GP32) (GP35); Envelope glycoprotein E2
(GP68) (GP70) (NS1); Protein P7: Nonstructural protein NS2 (P21)
(EC 3.4.22.-); Protease/helicase NS3 (P70) (Hepacivirin)
(EC 3.4.22.-); Protease/helicase NS3 (P70) (Hepacivirin)
(EC 3.4.21.98); Nonstructural protein NSAR (P4); Nonstructural protein NSAB (P77); Nonstructural protein NSSB (P66) (P70) (RNA-directed RNA polymerase) (EC 2.7.7.48)].
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             9; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      [RNA](N).
-1-SUBUNT: THE VIRION OF THIS VIRUS IS A NUCLECCAPSID COVERED BY
LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
PROTEIN M AND GLYCOPROTEIN E. THE NUCLECCAPSID IS A COMPLEX OF
PROTEIN C AND MRNA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY S29.
                                                                                                                                                                                                                                                                                                                                                                                        PRT; 3010 AA
10; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                            174 STRGVAKAVDFIPVESLETTMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      EMBL; M84754; -; NOT_ANNOTATED_CDS
                                                                                                                                                                                                                                                                                                                                                                                        STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       PDB; 1N64; 25-FEB-03
PDB; 1NS3; 08-APR-98
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         PIR; A40244; GNWVTW
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               WCBI_TaxID=31645;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Hepacivirus
                                                                                                                                                                                                                                                                                                                                                                                      POLG_HCVTW
P29846;
                                                                       셤
                                                                                                                                              ದ್ದ
                                                                                                                                                                                      õ
                                                                                                                                                                                                                      q
                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                        õ
                                                                                                             å
```

```
CORE PROTEIN (POTENTIAL).
MATRIX PROTEIN (POTENTIAL).
MAJOR ENVELOPE PROTEIN E (POTENTIAL).
NONSTRUCTURAL PROTEIN NSI/E2 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS2 (POTENTIAL).
PROTEASE/HELICASE NS3 (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         NONSTRUCTURAL PROTEIN NS4A (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4B (POTENTIAL).
RNA-DIRECTED RNA POLYMERASE (POTENTIAL)
POTENTIAL.
                                                                                                                                                                                                                                                                                         SMART; SM00487; DEXDC; 1.
Polyprotein; Glycoprotein; Transferase; RNA-directed RNA polymerase;
Core protein; Coat protein; Envelope protein; Helicase; ATP-binding;
Transmembrane; Nonstructural protein; Hydrolase; Serine protease;
                                                                                                                                                                                                                                                                                                                                                                     REMOVED FROM CAPSID PROTEIN C BY THE CELLULAR AMINOPEPTIDASE.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SIMILARITY).
SIMILARITY).
SIMILARITY).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              (POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           (POTENTIAL POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Length 3010;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        MW; AAD267D55CDFE215 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        (BY
(BY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CHARGE RELAY SYSTEM
CHARGE RELAY SYSTEM
CHARGE RELAY SYSTEM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       DB 1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ATP (POTENTIAL).
DECH BOX.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       8e-70;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       82.4%; Score 837.5; 78.4%; Pred. No. 8e-7 tive 18; Mismatches
                                                         erpro; IPRO07095; RN_pol_DS_PS.

m; PF01543; HCV_capsid: 1.

m; PF01543; HCV_capsid: 1.

m; PF01543; HCV_capsid: 1.

m; PF01560; HCV_cors; 1.

m; PF01560; HCV_NS1; 1.

m; PF01539; HCV_NS2; 1.

m; PF01006; HCV_NS3; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                N-LINKED
N-LINKED
N-LINKED
N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  N-LINKED
N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             N-LINKED
N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             N-LINKED
                                                                                                                                                                                                Pfam; PF01006; HCV_NS4a; 1.
Pfam; PF01001; HCV_NS4b; 1.
Pfam; PF01506; HCV_NS5a; 1.
Pfam; PF00271; helicasec: 1.
Pfam; PF00998; Viral_RGRP; 1.
ProDom; PD186062; HCV_NS1; 1.
                                                HCV_NS5a.
HCV_RdRP.
                              HCV_NS4b.
                 HCV_NS4a
 HCV_NS3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          AA; 327047
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               645
2041
2077
2240
2529
2788
IPR004109;
IPR000745;
                                                                              InterPro; IPR007095;
                              IPR001490;
                                                InterPro; IPR002868;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Similarity
                                                                                                                                                                                 Pfam; PF02907; H
Pfam; PF01006; H
Pfam; PF01001; H
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   160;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Query Match
Best Local
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Matches
```

;;

InterPro; IPR001410; DEAD.

MEROPS; S29.001; -.

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bloinformatics and the EMEL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a licensee agreement (See http://www.isb-sib.ch/announce/or send an email to licenseelabs.ch.ch.
                                                                                                                                                                                               1005 RRGREILLGPADSLEGRGWRLLAPITAYAQQTRGLFGCIITSLTGRDKNQVEGEVQVVST 1064
                                                           ATQTFLATSINGVLWTVYHGAGTRTIASPKGPVTQMYTNVDKDLVGWQAPQGSRSLTPCT 113
                                                                                                                                                                CGSSDLYLVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAV 173
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          01-APR-1993 (Rel. 25, Created)
01-APR-1993 (Rel. 25, Last sequence update)
15-SEP-2003 (Rel. 25, Last sequence update)
64-67-7003 (Rel. 42, Last annotation update)
65-7004 (Rel. 42, Last annotation update)
65-7005 (Rel. 42, Last annotation update)
67-707 (MRI); Protein P1 (GP83) (GP35); Envelope glycoprotein E2 (GP83) (GP35) (GP35); Envelope glycoprotein E2 (GP83) (GP35); Envelope glycoprotein E2 (GP83) (GP36) (Rel. 42.2-1); Protein P7; Nonstructural protein NS4A (P4); Nonstructural protein NS4B (P27); Nonstructural protein NS5A (P56); Nonstructural protein NS5A (P65); Nonstruc
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  (RNA)(N).
SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          LIPOPROTEIN ENVELOPE, THE ENVELOPE CONSISTS OF TWO PROTEINS: PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MRNA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY S29.
                                                                                                                                                                                                                                                                                                  174 STRGVAKAVDFIPVESLETTMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              MEDLINE-92295714; PubMed-1318627;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  EMBL; D11168; BAA01943.1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           PIR; A45573; A45573.
PDB; IA1Q; 25-MAR-98.
PDB; IJXP; I4-JAN-98.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  NCBI_TaxID-31642;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                POLG_HCVJT
                                                                                                                                                                114
                                                                                                                                                                                                                셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                             줨
                                                                                                     g
                                                                                                                                                                                                                                                                      ð
                                                     ò
                                                                                                                                                           ô
```

```
CELLULAR AMINOPEPTIDASE.
CAPEID PROTEIN (COTEWITAL).
MAJER ENVELOPE PROTEIN E (POTEWITAL).
MAJOR ENVELOPE PROTEIN E (POTEWITAL).
NONSTRUCTURAL PROTEIN NAI,F2 (POTEWITAL).
PROTEASE/HELICASE NS3 (POTEWITAL).
NONSTRUCTURAL PROTEIN NS4 ROTEWITAL).
NONSTRUCTURAL PROTEIN NS48 (POTEWITAL).
RNA-DIRECTED RNA POLYMERASE (POTEWITAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Gaps
                                                                                                                                                                                                                                                                                       Polyprotein, Glycoprotein, Transferase, RNA-directed RNA polymerase, Core protein, Coat protein; Envelope protein; Helicase; ATP-binding, Transmembrane, Nonstructural protein; Hydrolase; Serine protease; 3D-structure.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SIMILARITY).
SIMILARITY).
SIMILARITY).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Score 827.5; DB 1; Length 3010;
Pred. No. 6.9e-69;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          (POTENTI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              (POTENT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          18; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            94A1C77435D642BB CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                 POTENTIAL.
CHARGE RELAY SYSTEM (CHARGE RELAY SYSTEM)
ATP (POTENTIAL).
DECH BOX.
N-LINKED (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Mismatches
                                                                   terPro; IPR002868; HCV_NS4b.
terPro; IPR002166; HCV_RGRP.
terpro; IPR007095; RNA_CO_DS_FS.
"EPPCO; IPR007094; RNA_CO_DS_FS.
"PP01543; HCV_caps1d; 1.
"I) PP01542; HCV_corp."
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        20;
                                                                                                                                                                                                                                PF01506; HCV_NS5a; 1.

PF00271; helicase_C; 1.

PF00998; Viral_RGRP; 1.

; PD186062; HCV_NS1; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    81,48;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          AA; 326573
                                                                                                                                                                                   HCV_NS2; 1.
HCV_NS3; 1.
HCV_NS4a; 1
HCV_NS4b; 1
                                                                                                                                                   HCV_core;
HCV_env; 1
HCV_NS1; 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Matches 157; Conservative
                                                                                                                                                                                                                                                                               ; DEXDC;
                                                       IPR004109;
IPR000745;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Best Local Similarity
                                                                                                                                                                                                 PF02907;
PF01006;
                                                                                                                                                                                                                                                                    ProDom; PD18606
SMART; SM00487;
                                                                                                                                                                                                                        PF01001;
                                                                                                                                                                                                                                                         Pfam; PF00998;
         InterPro;
InterPro;
InterPro;
InterPro;
InterPro;
InterPro;
InterPro;
InterPro;
                                                                                                                InterPro;
                                                                                                     InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ACT_SITE
NP_BIND
SITE
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ACT_SITE
ACT_SITE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                    FRANSMEM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                           CHAIN
CHAIN
CHAIN
                                                                                                                                                                                                          Pfam;
                                                                                                                                                                                                                                                                                                                                                                                   CHAIN
                                                                                                                                        Pfam;
                                                                                                                                                                                                                                                                                                                                                             CHAIN
                                                                                                                                                                                                                        Pfam;
                                                                                                                                                                                                                                               Pfam;
                                                                                                                                                                                                                                                                                                                                                                          CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                           CHAIN
```

3 KKGSVVIVGRIN-----LSGDTAYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVST

ö

InterPro; IPR001410; DEAD.

MEROPS; S29.001; -.

MEROPS; U39.001; -

```
114 CGSSDLYLVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAV 173
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        MEDLINE-98277846; PubMed-958891;
A Pan Y., Li Y., Munshi S., Sardana V., Cole J.L., Sardana M.,
A Steinkuchler C., Tomel L., de Francesco R., Kuo L.C., Chen Z.;
Complex of NS3 protease and NS4A peptide of BK strain hepatitis C
virus: a 2.2-A resolution structure in a hexagonal crystal form.*;
IL Protein Sci. 7:87-847(1998).
- FUNCTION: THE SMALL PROTEINS NS2A, NS2B, NS4A AND NS4B ARE
HYDROPHOBIC, SUGGESTING A POSSIBLE MEMBRANE RELATED FUNCTION.
NS3 AND NS5 MAY PLAY A ROLE IN THE VIRAL RNA REPLICATION.
- CATALTIC ACTIVITY: Hydrolysis of four peptide boads in the viral precursor polyprotein, commonly with Asp or Glu in the P6
position, Cys or Thr in Pl and Ser or Ala in Pl.
- CATALTIC ACTIVITY: N nucleoside triphosphate +
                                                                                                                                                                                                                                                                                                                                                                                      01-AUG-1992 (Rel. 23, Created)
01-AUG-1992 (Rel. 23, Last sequence update)
15-SEP-2003 (Rel. 23, Last sequence update)
64-Rel. 2003 (Rel. 42, Last annoration update)
65-Rel. 2003 (Rel. 42, Last annoration update)
66-Rel. 2004 (Rel. 42, Last annoration update)
67-Rel. 42, Last annoration protein C (Core protein) (P22);
67-Rel. (GPR8) (GP70) (NG1); Protein E1 (GP83) (GP35); Envelope glycoprotein E2 (GP88) (GP70) (NG1); Protein R7; Nonstructural protein NS4 (Rel. 42, 2-); Protease/helicase NS3 (P70) (Hepacivitin)
6C 3.4.21.98); Nonstructural protein NS4 (P4); Nonstructural protein NS4B (P65); NONSTructural protein NS4B (P65) (P70) (Rel. 40); Nonstructural protein NS4B (P65) (P70) (RF) (P70) (P70)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     X-RAY CRYSTALLOGRAPHY (2.4 ANGSTROMS) OF 1027-1215.
MEDLINE-97015088; pubMed-8861916;
LOVE R.A., Parge H.E., Mickersham J.A., Hostomsky Z., Habuka N.,
MOOMAW E.W., Adachl T., Hostomska Z.;
"The crystal structure of hepatitis C virus NS3 proteinase reveals a trypsin-like fold and a structural zinc binding site.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   the viral
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Takamizawa A., Mori C., Fuke I., Manabe S., Murakami S., Fujita J., Onishi E., Andoh T., Yoshida I., Okayama H.; "Structure and organization of the hepatitis C virus genome isolated
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS) OF 1027-1210 AND 1678-1691
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          J. Virol. vo.--.

[2]
SEQUENCE OF 1487-1500.
MEDLINE-96235224; PubMed-8647104;
Borowski P., Helland M., Ochlmann K., Becker B., Korneteky L.;
"Non-structural protein 3 of hepatitis C virus inhibits
"Non-structural protein 3 of hepatitis C virus inhibits
phosphorylation mediated by cAMP-dependent protein kinase.";
T Riochem. 237:611-618(1996).
                                                                                                                                                                                                                                                                                                                                                       PRT; 3010 AA.
                                                                                                                                                                                                                                        174 STRGVAKAVDFIPVESLETTMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SEQUENCE FROM N.A.
MEDLINE-91140698; PubMed~1847440;
                                                                                                                                                                                                                                                                                                                                                       STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Cell 87:331-342(1996)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          from human carriers.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Hepacivirus.
NCBI_TaxID=11105;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (RNA) (N)
                                                                                                                                                                                                                                                                                                                                                       POLG_HCVBK
P26663;
                                                                                                                                                                                                                                                                                                                                       POLG_HCVBK
                                                                                                                                                                                                                                                                                                                                                          q
                                                                                                                                                                                                                      à
                                                                                                                                                                                                                                                           셤
                                                                                                a
                                                                                                                                     ò
                                                           ò
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMED outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license also removed. Gsee http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CELLULAR AMINOPEPTIDASE.
CAPSID PROTEIN ( POTENTIAL).
MATRIX PROTEIN (POTENTIAL).
MAJOR ENVELOPE PROTEIN E (POTENTIAL).
NONSTRUCTURAL PROTEIN NSI/E2 (POTENTIAL).
NONSTRUCTURAL PROTEIN NSI/E2 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4A (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4A (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               RNA-DIRECTED RNA POLYMERASE (POTENTIAL).
POTENTIAL.
 NUCLEOCAPSID COVERED BY A
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Polyprotein; Glycoprotein; Transferase; RNA-directed RNA polymerase;
Core protein; Coat protein; Envelope protein; Helicase; ATP-binding;
Transmembrane; Monstructural protein; Hydrolase; Serine protease;
               LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS: PROTEIN M AND GLICOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MRNA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      REMOVED FROM CAPSID PROTEIN C BY
                                                                      SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY S29.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CHARGE RELAY SYSTEM.
CHARGE RELAY SYSTEM.
CHARGE RELAY SYSTEM.
ATP (POTENTIAL).
DECH BOX.
SUBUNIT: THE VIRION OF THIS VIRUS IS A
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          InterPro; IPR007095; RNA_POL_DS_PS.
InterPro; IPR007094; RNA_POL_DS_PS.
Pfam; PF01542; HCV_capsid; 1.
Pfam; PF01539; HCV_core; 1.
Pfam; PF01539; HCV_env; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           HCV_capsid.
HCV_core.
HCV_env.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   HCV_NS2.
HCV_NS3.
HCV_NS4a.
HCV_NS4b.
HCV_NS5a.
HCV_NS5a.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Pfam; PF01538; HCV_NS2; 1. Pfam; PF02907; HCV_NS3; 1. Pfam; PF02006; HCV_NS4s; 1. Pfam; PF01001; HCV_NS4b; 1. Pfam; PF01506; HCV_NS5a; 1. Pfam; PF00998; Viral_RdRP; 1. Probom; PD186062; HCV_NS1; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    HCV_NS1.
                                                                                                                                                                                                                                                           EMBL; M58335; AAA72945.1; -
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         115
191
383
729
1006
1615
1862
2013
3010
369
1103
11165
11165
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SMART; SM00487; DEXDC;
                                                                                                                                                                                                                                                                                              1A1Q, 25-MAR-98.
1JXP, 14-JAN-98.
1C2P, 15-NOV-00.
1C2J, 08-NOV-99.
1GX5, 09-APR-02.
1GX6, 10-APR-02.
1GX6, 26-JUN-00.
80HM; 20-APR-99.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                InterPro; IPR002531;
InterPro; IPR002518;
InterPro; IPR004109;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     InterPro; IPR000745;
InterPro; IPR001490;
InterPro; IPR002868;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            InterPro; IPR002166;
InterPro; IPR007095;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         InterPro; IPR001410;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            IPR002521;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             InterPro; IPR002522;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                IPR002519;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           116
192
384
730
1007
1007
11616
347
347
1108
11107
11165
11165
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  MEROPS; S29.001
MEROPS; U39.001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  3D-structure.
                                                                                                                                                                                                                                                                                   A38465;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ACT_SITE
ACT_SITE
ACT_SITE
NP_BIND
SITE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  TRANSMEM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CHAIN
CHAIN
CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CHAIN
                                                                                                                                                                                                                                                                                                                    PDB;
PDB;
PDB;
PDB;
PDB;
PDB;
```

```
1064
                                                                                                                                                                                                                                                                                                                                                                  3 KKGSVVIVGRIN-----LSGDTAYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVST
                                                                                                                                                                                                                                                                                                                                 Gaps
                      (POTENTIAL)
                                                                                          (POTENTIAL).
(POTENTIAL).
(POTENTIAL).
(POTENTIAL).
           (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                 ő
                                                                                                                                                                                                                                                                                                                     Length 3010;
                                                                                                                                                                                                                                                                                                                                 Indels
                                                                                                                                                                                                                                                                                                         327189 MW; F8422D5ECCFDFD9C CRC64;
                                                                                                                                                                                                                                                                                                                     DB 1;
                                                                                                                                                                                                                                                                                                                                 18;
                                                                                                                                                                                                                                                                                                                    81.0%; Score 823.5; DB 1 76.5%; Pred. No. 1.6e-68;
                                                                         (GLCNAC.
(GLCNAC.
(GLCNAC.
   GLCNAC
(GLCNAC
(GLCNAC
(GLCNAC
(GLCNAC
(GLCNAC
(GLCNAC
(GLCNAC
(GLCNAC
(GLCNAC
(GLCNAC)
                                                                                          (GLCNAC.
(GLCNAC.
(GLCNAC.
                                                                                                                                                                                                                                                                                                                                21; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                      N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                 STRGVAKAVDFIPVESLETTMRSP 197
                                                                                                                                                                                                                                                                                                                                Conservative
                                                                               1157
1161
1163
1166
                                                                                                                                                                                                                                                                                                   1688
                                                                                                                                                                                                                                                           1168 117
1172 111
11175 1118
11187 1118
11198 1120
11203 120
11680 166
                                                                                                                                                                                                                                                                                                                          Similarity
                                                                                                                                                                                       1101
1104
1120
1120
1122
1129
1139
                                                                                                                                                                                                                                                1162
                                                                                                                                                         082
086
090
093
093
                                                                                                                                                                                                                                                                                                                                Matches 156;
                                                                                                                                                                                                                                                                                                   STRAND
                                                                         CARBOHYD
CARBOHYD
CARBOHYD
                                                                                          CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                 174
   CARBOHYD
                                                                                                                                                                                                                                                                                                                     Query Match
Best Local 8
                                                                                                                                                                     STRAND
TURN
STRAND
STRAND
                                                                                                                                                                                                  STRAND
STRAND
STRAND
STRAND
STRAND
STRAND
HELLX
TURN
STRAND
                                                                                                                       HELIX
STRAND
STRAND
STRAND
                                                                                                                                                    STRAND
HELIX
TURN
                                                                                                                                                                                                                                                                                  STRAND
HELIX
TURN
                                                                                                                  STRAND
à
                                                                                                                                                                                                                                                                                                                                                       셤
                                                                                                                                                                                                                                                                                                                                                                   ç
                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                 ò
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMED outstation the Buropean Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
        {RNA}(N).
SUBDINT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MRNA.
                                                                                                                                                               SEQUENCE FROM N.A.
MEDIJUR-91088550; PubMed-2175903;
Kato N., Hilkata M., Ootsuyama Y., Nakagawa M., Ohkoshi S.,
Sugimura T., Shimotohno K.;
"Molecular cloning of the human hepatitis C virus genome from
"Molecular cloning of the human hepatitis P.,
Proc. Natl. Acad. Sci. U.S.A. 87:9524-9528(1990).
                                                                                                                                                                                                                                                                                                                                                                                                                     SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY S29
PRT; 3010 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 HCV_capsid.
HCV_core.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     HCV_NS4b.
HCV_NS5a.
HCV_RdRP.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     HCV_env.
HCV_NS1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            HCV_NS48
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        HCV_NS2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       EMBL; D90208; BAA14233.1; -.
STANDARD;
                                                                                                                                                                                                                                             DISCUSSION OF SEQUENCE.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          IPR002521;
IPR002519;
IPR002531;
IPR002538;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        IPR001410;
IPR002522;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 PIR; A39253; GNWVCJ.
HSSP; P26663; LJXP.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          IPR002166;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    MEROPS; S29.001; -. MEROPS; U39.001; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  IPR004109;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            IPR000745;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      IPR001490;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              IPR002868;
                                                                                                                                             NCBI_TaxID-11116;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              InterPro;
InterPro;
POLG_HCVJA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              InterPro;
InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            InterPro;
```

œ

```
1005 RRGKEILLGPADSFGEQGWRLLAPITAYSQQTRGLLGCIITSLTGRDKNOVDGEVQVLST 1064
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ATQTFLATSINGVLWTVYHGAGTRTIASPKGPVTQMYTNVDKDLVGWQAPQGSRSLTPCT 113
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    114 CGSSDLYLVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAV 173
                                                                                                                                                  3 KKGSVVIVGRINLSGD-----TAYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVST
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      9; Gaps
                                                                                                                                                                                                                                                                                             SIMILARITY).
SIMILARITY).
SIMILARITY).
                                                                                                                                                                                                                                                                                                                                                                                                                                       (POTENTIAL).
(POTENTIAL).
(POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                            (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                             (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                          (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                               (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    DB 1; Length 3010;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Indels
                                                                                                                                                                                                                                                                                            RELAY SYSTEM (BY
RELAY SYSTEM (BY
RELAY SYSTEM (BY
                                                                                                                                                                                                                                                                                                                                                                                                            (GLCNAC. . .)
                                                                                                                                                                                                                                                                                                                                                                                                                                        (GLCNAC. . .)
                                                                                                                                                                                                                                                                                                                                                                                                                                (GLCNAC. . .)
                                                                                                                                                                                                                                                                                                                                                                                 (GLCNAC. . .)
                                                                                                                                                                                                                                                                                                                                                                                                                      (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         81.0%; Score 823.5; DB 1;
75.5%; Pred. No. 1.6e-68;
1ive 23; Mismatches 18;
                                                                                                                                                                                                                                                                                                                                                                                                                                      N-LINKED GLCNAC.
                                                                                                                                                                                                                                                                                                                                          N-LINKED (GLCNAC.
N-LINKED GLCNAC.
                                                                                                                                                                                                                                                                                                                                                              (GLCNAC.
                                                                                                                                                                                                                                                                                                                          ATP (POTENTIAL).
DECH BOX.
                                                                                                                                                                                                                                                                                     POTENTIAL.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           STRGVAKAVDFIPVESLETIMRSP 197
                                                                                                                                                                                                                                                                                               CHARGE
                                                                                                                                                                                                                                                                                                                 CHARGE
InterPro; IPR01650; Helicase_C.
InterPro; IPR07095; RNA_pol_DS_PS.
InterPro; IPR007094; RNA_pol_PSvir.
                           Pfam; PF01543; HCV_capsid; 1. Pfam; PF01542; HCV_capsid; 1. Pfam; PF01542; HCV_capsid; 1. Pfam; PF01542; HCV_capvid; 1. Pfam; PF01560; HCV_NS1; 1. Pfam; PF01560; HCV_NS2; 1. Pfam; PF01006; HCV_NS3; 1. Pfam; PF01001; HCV_NS4; 1. Pfam; PF01001; HCV_NS4; 1. Pfam; PF001506; HCV_NS5a; 1. Pfam; PF00999; Viral_RGRP; 1. Probon; PD186062; HCV_NS1; 1. SWART; SW00487; DEXDC; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  327017
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Query Match
Best Local Similarity 75.5
Matches 154; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               2240
2788
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ž
                                                                                                                                                                                                                                                                           CHAIN
TRANSMEM
ACT_SITE
ACT_SITE
ACT_SITE
NP_BIND
                                                                                                                                                                                                                                                                                                                                                               CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                          CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                   CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                             CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                        CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CARBOHYD
                                                                                                                                                                                                                                                                                                                                             CARBOHYD
                                                                                                                                                                                                                                                                                                                                                       CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                  CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                         CARBOHYD
   셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
```

1185 CTRGVAKAVDFIPVESMETTHRSP 1208

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMED. Custration the Buropean Bioinformatics institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@igb-sib.ch).
                                                          01-A0G-1992 (Rel. 23, Created)
01-A0G-1992 (Rel. 23, Last sequence update)
01-A0G-1992 (Rel. 24). Last sequence update)
01-A0G-1992 (Rel. 41, Last amnotation update)
Genome polyprotein (Contains Capsid protein C (Core protein) (P22);
Envelope glycoprotein El (GP32) (GP35); Envelope glycoprotein E2
(GP68) (GP70) (NS1); Protein P7; Nonstructural protein NS2 (P21)
(EC 3.4.22.-); Protease/helicase NS3 (P70) (Hepacivirin)
(EC 3.4.21.98); Nonstructural protein NS5A (P4); Nonstructural protein
NS5B (P27); Nonstructural protein NS5A (P56); Nonstructural protein
NS5B (P66) (P70) (RMa-directed RNA pollymerase) (EC 2.7.7.48)].
Hepatitis C virus (isolate HC-J8) (HCV).
                                                                                                                                                                                                                                                                                                                                                                                                                             (RNA)(N).
SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MRNA.
                                                                                                                                                                                                                                                                                                                            MEDLINE-9223022; PubMed-1314459; Okada S.-I., Yamamoto K., Lizuka H., Tanaka T., Okamoto H., Kurai K., Okada S.-I., Yamamoto K., Lizuka H., Tanaka T., Fukuda S., Tauda F., Mishiro S.; Fukuda S., Tauda F., Figula Poor Figuli-length sequence of a hepetitis C virus genome having poor homology to reported isolates; comparative study of four distinct
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 .!- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY S29.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Interpro; IPR007095; RNA_pol_DS_PS.
Interpro; IPR007094; RNA_pol_PSvir.
Pfam; PF01543; HCV_capsid; 1.
Pfam; PF01542; HCV_core; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      HCV_NS4a.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        HCV_NS5a.
HCV_RdRP.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       HCV_NS3.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                EMBL; D10988; BAA01761.1; -.
                               STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     InterPro; IPR001410;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       InterPro; IPR002531;
InterPro; IPR002518;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        InterPro; IPR000745;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      InterPro; IPR001490;
InterPro; IPR002868;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          InterPro; IPR002166;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     InterPro; IPR002522;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       InterPro; IPR002519;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        InterPro; IPR004109;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    PIR; A40250; GNWVJB.
HSSP; P27958; IHEI.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  MEROPS; S29.001; -
                                                                                                                                                                                                                                                                                                                SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                NCBI_TaxID-11115;
                               POLG_HCVJ8
                                                                                                                                                                                                                                                                                                                                                                                                                   genotypes.
RESULT 7
POLG_HCVJ8
```

```
fam:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         fam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Pfam:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                79 IASPKGPVTQMYTNVDKDLVGWQAPQGSRSLTPCTCGSSDLYLVTRHADV1PVRRRGDSR 138
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            19 TAYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVSTATQTFLATSINGVLMTVYHGAGTRT 78
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    139 GSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVSTRGVAKAVDFIPVESLETIMRSP 197
                                                                                                                                                        CELLULAR AMINOEPTIONSE.
CAPSID PROTEIN C (POTENTIAL).
MATRIX PROTEIN C (POTENTIAL).
MAJOR ENVELOPE PROTEIN E (POTENTIAL).
NONSTRUCTURAL PROTEIN NS! (POTENTIAL).
NONSTRUCTURAL PROTEIN NS2 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS48 (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       0; Gaps
                                                                                                               Polyprotein; Glycoprotein; Transferase; RNA-directed RNA polymerase; Core protein; Coat protein; Bnvelope protein; Helicase; APP-binding; Transmembrane; Nonstructural protein; Hydrolase; Serine protease.

INIT_MET
                                                                                                                                                                                                                                                                                  CHARGE RELAY SYSTEM (BY SIMILARITY).
CHARGE RELAY SYSTEM (BY SIMILARITY).
CHARGE RELAY SYSTEM (BY SIMILARITY).
                                                                                                                                                                                                                                                                                                                                                                                                    (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                              (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                          (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                      (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                            (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ( POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Length 3033;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    30; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1A173E7E3381FD1A CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               , DB 1;
1.3e-54;
                                                                                                                                                                                                                                                                                                                                                                                                                                            N-LINKED GLCNAC
                                                                                                                                                                                                                                                                                                                                                     (GLCNAC.
(GLCNAC.
(GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                    (GLCNAC.
(GLCNAC.
(GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                   (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                        (GLCNAC
                                                                                                                                                                                                                                                                                                                                             N-LINKED (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                       (GLCNAC
                                                                                                                                                                                                                                                                                                                      ATP (POTENTIAL)
DECH BOX.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     24; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                66.4%; Score 675;
                                                                                                                                                                                                                                                                                                                                                                                      N-LINKED (
N-LINKED (
N-LINKED (
                                                                                                                                                                                                                                                                                                                                                                                                                        N-LINKED (
                                                                                                                                                                                                                                                                                                                                                     N-LINKED
N-LINKED
N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Pred. No.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         330177 MW;
Pfam: PF01539; HCV_env: 1.
Pfam: PP01560; HCV_NS1: 1.
Pfam: PF01538; HCV_NS2: 1.
Pfam: PF02907; HCV_NS3: 1.
Pfam: PF01006; HCV_NS4s: 1.
Pfam: PF01001; HCV_NS4b; 1.
Pfam: PF01506; HCV_NS5b; 1.
Pfam: PF01998; VITAL_RGRP: 1.
ProDom: PD186062; HCV_NS1: 1.
SWART; SN00487; DEXDC: 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          69.88;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1091
2038
2359
2811
                                                                                                                                                                                                                             1619
1866
2017
3033
                                                                                                                                                                                                                                                                                               1111
1169
1241
1323
196
209
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Similarity
                           Loun; PP02907; E. Pfam; PF01006; H. Pfam; PF01001; H. Pfam; PF01001; H. Pfam; PF0150f
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     125;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1034
                                                                                                                                                                                                                                                                                    ACT_SITE
ACT_SITE
ACT_SITE
NP_BIND
                                                                                                                                                                                                                                                              CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Query Match
                                                                                                                                                                                                                                                                                                                                                                    CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                               CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                        CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                    CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                               CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                          CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                      CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                           CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SEQUENCE
                                                                                                                                                                                                                                                                                                                                             CARBOHYD
                                                                                                                                                                                                                                                                                                                                                        CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Best Local
                                                                                                                                                                        CHAIN
                                                                                                                                                                                   CHAIN
CHAIN
CHAIN
CHAIN
CHAIN
CHAIN
   ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         셤
```

PRT; 3033 AA

STANDARD;

POLG\_HCVJ6 P26660;

200

RESULT 8 POLG\_HCVJ6

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation he European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/
                                                                                                                                                                                                                                                                                                                                                                   JULIANE-920444(1) PubMed-1658196;

AEDLINE-920444(1) PubMed-1658196;

AREDLINE-920444(1) PubMed-1658196;

AREDLINE-920444(1) PubMed-1658196;

ARCHIGA A., Miyakawa Y., Mayuan M.;

Machida A., Miyakawa Y., Miyaka M.;

Machida A., Miyaka M.;

Machida A., Miyakawa Y., Miyaka M.;

Machida A., Miyakawa Y., Miyaka M.;

Machida A., Miyakawa Y., Miyaka M.;

Machida A., Miyaka M.;

Machida A., Miyakawa Y., Miyaka M.;

Machida A., Miyakawa Y., Miyaka M.;

Machida A., Miyakawa M.;

Machida A., Miyakawa M.;

Machida A., Miyakawa M.;

Machida M.;

Machida A., Miyaka M.;

Machida M.;

Machi
01-AUG-1992 (NEL. 23, Last sequence update)
28 FEB-2003 (Rel. 21, Last sequence update)
Genome polyprotein [Contains: Capsid protein C (Core protein) (P22);
Envelope diycoprotein E1 (GP32) (GP35); Envelope glycoprotein E2
ENVELOPE G1900 (NS1); Protein P7; Nonstructural protein NS2 (P21)
(EC 3.4.22.-); Protease/helicase NS3 (P70) (Hepacivirin)
(EC 3.4.1.08); Nonstructural protein NSA (P4); Nonstructural protein
NSSB (P65); (P70) (RNA-directed RNA polymerase) (EC 2.7.7.48);
Nivuses; SSRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   [RNA](N).
-1- SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A LIPOPROTEIN ENVELOPE CONSISTS OF TWO PROTEINS: PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MENA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           -! - SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY 529
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              HCV_NS5a.
HCV_RdRP.
Helicase_C.
RNA_pol_DS_PS.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             HCV_capsid.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 HCV_core
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      HCV_NS4a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               HCV_NS4b
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             EMBL; D00944; BAA00792.1; -. PIR; JQ1303. HSSP; P27958; 1HEI. MEROPS; S29.001; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                HCV_NS1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       HCV_NS2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     PF01543; HCV_capsid;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    HCV_env; 1.
HCV_NS1; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   HCV_NS2; 1
HCV_NS3; 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                HCV_core;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   HCV_NS4a;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      HCV_NS4b;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  HCV_NS5a;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    IPR002518;
IPR004109;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  IPR001650;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      IPR007095;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       InterPro; IPR001410;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           InterPro; IPR002522;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          IPR002519;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   IPR000745;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               IPR001490;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                IPR002868;
                                                                                                                                                                                                                                                                                                                                                      SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                       NCBI_TaxID-11113;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      PF01001;
PF01506;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              MEROPS; U39.003
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       InterPro;
InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                InterPro;
InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        interPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 nterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Pfam;
```

```
79 IASPKGPVTQMYTNVDKDLVGWQAPQGSRSLIPCTCGSSDLYLVTRHADVIPVRRRGDSR 138
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                78
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                139 GSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVSTRGVAKAVDFIPVESLETIMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              TAYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVSTATQTFLATSINGVLWTVYHGAGTKT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Arabidopsis thaliana (Mouse-ear cress).
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
                                                                                                                                                  PROTEASE/HELICASE NS3 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4A (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4B (POTENTIAL).
RNA-DIRECTED RNA POLYMERASE (POTENTIAL).
                                SMART: SM00487; DEXDC: 1.

Polyprotein; Glycoprotein; Transferase; RNA-directed RNA polymerase; Core protein; Glycoprotein; Envelope protein; Heliasae; Arp-binding; Transmenbrane; Nonstructural protein; Hydrolase; Serine protease.

INIT_MET
                                                                                             CAPSID PROTEIN C POTENTIAL).
MATRIX PROTEIN (POTENTIAL).
MAJOR ENVELOPE PROTEIN E (POTENTIAL).
NONSTRUCTURAL PROTEIN NS2 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS2 (POTENTIAL).
                                                                                                                                                                                                      SIMILARITY).
SIMILARITY).
SIMILARITY).
                                                                                                                                                                                                                                                                                          (POTENTIAL).
(POTENTIAL).
(POTENTIAL).
(POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                          (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                   (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                             (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                  (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                 (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ö
                                                                                                                                                                                                                                                                                                                                                                                                            (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                          (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                      66.3%; Score 674; DB 1; Length 3033; 68.7%; Pred. No. 1.6e-54; ive 27; Mismatches 29; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                  F957F5C1A273BE9E CRC64;
                                                                                                                                                                                                      CHARGE RELAY SYSTEM (BY
CHARGE RELAY SYSTEM (BY
CHARGE RELAY SYSTEM (BY
                                                                                      CELLULAR AMINOPEPTIDASE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (GLCNAC. .
                                                                                                                                                                                                                                                                               (GLCNAC.
(GLCNAC.
(GLCNAC.
                                                                                                                                                                                                                                                                                                                        (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                      (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                      (GLCNAC.
                                                                                                                                                                                                                                       (POTENTIAL).
                                                                                                                                                                                                                                                             (GLCNAC
                                                                                                                                                                                                                                                                                                                                              (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                            (GLCNAC
                                                                                                                                                                                                                                                                      GLCNAC
                                                                                                                                                                                                                                                                                                              (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                           (GICNAC
                                                                                                                                                                                             POTENTIAL.
                                                                                                                                                                                                                                                           N-LINED
                                                                                                                                                                                                                                                DECH BOX.
Pfam; PF00271; helicase_C; l. Pfam; PF00998; Viral_RdRP; l. ProDom; PD186062; HCV_NS1; l.
                                                                                                                                                                                                                                                                                                                                                                                                                                                  329165
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Conservative
                                                                                                                                                                                                                                                                                                                                                                                                          649
1091
2038
2811
                                                                                                                                                                                                                                                                                                                                                                                                                                                  3033 A.A.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Local Similarity
tes 123; Conserv
                                                                                                                                                                                                                           1169
1234
1320
196
209
                                                                                                                                                                                                     ACT_SITE
                                                                                                                                                                                                                           ACT_SITE
NP_BIND
SITE
                                                                                                                                                                                                                                                                                                                                                                                                          CARBOHYD
CARBOHYD
                                                                                                                                                  CHAIN
CHAIN
CHAIN
CHAIN
TRANSMEM
                                                                                                                                                                                                                                                                                                                       CARBOHYD
                                                                                                                                                                                                                                                                                                                                   CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                              CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Query Match
                                                                                                                                                                                                                                                           CARBOHYD
                                                                                                                                                                                                                                                                                           CARBOHYD
                                                                                                                                                                                                                                                                                                    CARBOHYD
                                                                                                                                                                                                                                                                                                              CARBOHYD
                                                                                                                                                                                                                                                                                                                                             CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                  CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                     CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                               CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                  SEQUENCE
                                                                                                                                                                                                                                                                                CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                           CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 HHOA ARATH
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     RESULT 9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ŏ
```

```
WEBLIAND-1702;

WEBLIAND-1702;
```

```
modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      12;
   as its content is in no way
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           95 KTSPSVVYIEAIELPKTSSGDIL------ 127
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ----SLLSPRPISYLK------GSSGGPLLCPAGHAVGIFRAAVSTR--GVAKAV 182
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SGLGREIPSPNGKSISEAIQTDADINSGNSGGPLLDSYGHTIGVNTATFTRKGSGMSSGV 294
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       182 --DPDNDLAVLKIETEGRELNPVVLGTSNDLRVGSCFAI----GNPYGYENTLTIGVV 234
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             2 KKKGSVVIVGRINL---SGDIAYAQQIRGEEGCQEISQIGRDKNQVEGEVQIVSTAIQI 57
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           58 FLATSINGVLW-----TVYH--------GAGTRTIASPKGPVTQM 89
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      90 YINVDKDLVGWQAPQGSRSLTPCTCGSSDLYLVTRHADVIPVRRRGDSRG------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        STRAIN-AICC 15692 / PAO1;
STRAIN-AICC 15692 / PAO1;
MEDLINE-20437337; PubMed-10984043;
MEDLINE-20437337; PubMed-10984043;
MEDLINE-20437337; PubMed-10984043;
Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warrener P., Hickey M.J., Brinkman P.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M. Garber R.L., Goltry L., Tolentino E., Westbrock-Wadman S., Yuan Y., Brody L.L., Coulter S.N., Folger K.R., Kas A., Larbig K., Lim R.M., Snith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T., Reizer J., Saler H.H., Hancock R.E.W., Lory S., Olson M.V.; "Complete genome sequence of Pseudomonas aeruginosa PAO1, an
                                                                                                                                                                                                                                                                                                                                                                                                                      Hydrolase: Serine protease: Chloroplast; Thylakoid; Transit peptide.
TRANSIT 1 26 CHLOROPLAST (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Pseudomonas aeruginosa.
Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
Pseudomonadaceae; Pseudomonas.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   68; Indels 100;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CHARGE RELAY SYSTEM (POTENTIAL).
CHARGE RELAY SYSTEM (POTENTIAL).
CHARGE RELAY SYSTEM (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Match 8.6%; Score 87; DB 1; Length 321; Local Similarity 21.3%; Pred. No. 0.78; es 54; Conservative 31; Mismatches 68; Indels 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   -> G (IN REF. 1).
68DB81E0BD27A7A7 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    16-0CT-2001 (Rel. 40, Created)
116-0CT-2001 (Rel. 40, Last sequence update)
116-0CT-2001 (Rel. 40, Last annotation update)
Probable aromatic acid decarboxylase (EC 4.1.1.-).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              209 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         PROTEASE HHOA.
non-profit institutions as long
                                                                                                                                                                                                                                                     MEROPS; SOI.279; ...
InterPro; IPR001940; Protease2C.
InterPro; IPR001254; Ser_protease_Try.
Pfam: PF00089; trypsin; 1.
PRINTS; PR00834; PROTEASFS2C.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          THYLAKOID
                                                                                                                                                             EMBL; AF114386; AAF24060.1; -.
EMBL; AL021710; CAA16717.1; ALT_SEQ.
EMBL; AL161548; CAB78839.1; ALT_SEQ.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               321 AA; 34691 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                DF-IPVESLETTM 194
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        NCBI_TaxID-287;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              PAAD_PSEAE
Q9HX08;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ACT_SITE
CONFLICT
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 140
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  235
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   183
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ACT_SITE
ACT_SITE
                                                                                                                                                                                                                                                                                                                                                                                                                                                           TRANSIT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Query Match
Best Local
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           PAAD_PSEAE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SOUTH THE TENT TO SEE THE TENT THE TENT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               đ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation-the Buropean Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/
                                                                                                                                                                                                                                                                                                                                                                                                                                                                   11;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             138 RGSLLSPR--PIS-----TLKGSSGGPLLCPAGHAVGIFRAAVSTRGVAKAVDFIPVES 189
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   92 NVDKDLVGWQAPQGSRSLTP-----CTCGSSDL-----YLVTRHADVIPVRRGDS 137
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  81 OND-----WMAPPASGSSAPNAMVICPCSTGTLSAVATGACNNLIERAADVALKER---- 131
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  80
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         43 QVEGEVQ-IVSTATQTFLATSINGVL------WTVYHGAGTRTIASPKGPVTQMYT 91
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 DEGLARATH STANDARD; PRT; 437 AA.

022609; Q9LR85;
16-007-2001 (Rel. 40, Created)
16-007-2001 (Rel. 41, Last sequence update)
28-FEB-2003 (Rel. 41, Last annotation update)
Portease Do-like 1, chloroplast precursor (EC 3.4.21.-).
Protease Do-like 1, chloroplast precursor (EC 3.4.21.-).
DEGPL OR DEGP OR AT3627925 OR K16N12.18.
Arabidopsis thaliana (Mouse-ear cress).
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
eurosids II; Brassicales; Brassicaceae; Arabidopsis.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  29 QEEREVHFLISKAAQLVMATETDVALPAKPQAMQAFLTEYCGAAAGQI------RVFG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        STRAIN-CV. Columbia;

MEDLINE-20363099; PubMed-10907853;

Raneko T., Katoh T., Sato S., Nakamura A., Asamizu E., Tabata S.;

Kaneko T., Katoh T., Sato Rabidopsis thaliana chromosome 3. II .

"Structural analysis of Arabidopsis thaliana chromosome 3. II .

Sequence features of the 4,251,695 bp regions covered by 90 Pl, TAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SEQUENCE FROM N.A., AND CHARACTERIZATION.
MEDLINE-98175982; PubMed-9507020;
Itzhaki H., Naveh L., Lindahl M., Cook M., Adam Z.;
Identification and characterization of Degp, a serine protease associated with the luminal side of the thylakoid membrane.";
J. Biol. Chem. 273:7094-7098(1998).
           Nature 406:959-964(2000).
-I. SIMILARITY: BELONGS TO THE POLYPRENYL P-HYDROXYBENZOATE /
PHENYLACRYLIC ACID DECARBOXYLASES FAMILY.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                 55;
                                                                                                                                                                                                                                                                                                           InterPro: IPR003382; Flavoprotein.
Pfam: PF02441; Flavoprotein; 1.
Hypothetical protein: Lyase; Decarboxylase; Complete proteome.
SEQUENCE 209 AA; 22367 MW; 01FD081CC495D3F6 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                  Watch 8.4%; Score 85.5; DB 1; Length 209; Local Similarity 27.9%; Pred. No. 0.64; hes 51; Conservative 16; Mismatches 61; Indels 5
                                                                                                                                                                                                                             or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                       EMBL; AE004818; AAG07406.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            DNA Res. 7:217-221(2000).
                                                                                                                                                                                                                                                                                              PIR; H83144; H83144.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  [3]
SEQUENCE OF 104-118.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SECUENCE FROM N.A
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             190 LET 192
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     LNT 189
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       and BAC clones
                                                                                                                                                                                                                                                                                                                                                                                                                       Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    RESULT 11
ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
```

```
AAMP_HUMAN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  REPEAT
REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           RPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Matches
           à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            a
                                                                                                                                          This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the ENBL cutstation the European Bioinformatics institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (5ee http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    265
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          150 VPQGSGSGFVWDKQGHIVTNYHVIRGASDLRVTLADQTTFDAKVVGFDQDKDVAVLRIDA 209
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               103 PQGSRSLTPCTCGSSDLYLV------TRHADVIPVRRGDSRGSLLSPRPI 147
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            REGULATION: INHIBITED BY PHENYLMETHYLSULFONYL FLUORIDE AND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Gaps
      Kieselbach T., Bystedt H., Schroeder W.P.;
Submitted (JUL-2000) to the SWISS-PROT data bank.
-!- FUNCTION: SERINE PROTEASE THAT IS REQUIRED AT HIGH TEMPERATURE.
MAY BE INVOLVED IN THE DEGRADATION OF DAMAGED PROTEINS. IN VIVO,
CAN DEGRADE BETA-CASEIN.
                                                                                                                                                                                                                                                                                                                                                      Transit peptide; Chloroplast; Thylakoid
                                                                                                                                                                                                                                                                                                                                                                                                                                                HSPPSSQLSNST -> SSTLFLHSPPSSHL (IN REF
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       148 SYL------KGSSGGPLLCPAGHAVGIFRAAVSTRGVAKAVDF-IPVESL 190
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SUBCELLULAR LOCATION; BOUND TO LUMINAL SIDE OF THE THYLAKOID
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 55;
                                                                                                                                                                                                                                                                                                                                                                                                                  CHARGE RELAY SYSTEM (POTENTIAL).
CHARGE RELAY SYSTEM (POTENTIAL).
CHARGE RELAY SYSTEM (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          8.2%; Score 83.5; DB 1; Length 437; 26.2%; Pred. No. 2.4;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                   V -> 1 (IN REF. 2).
P -> 5 (IN REF. 2).
G -> R (IN REF. 2).
G -> D (IN REF. 2).
LL -> HF (IN REF. 2).
L -> V (IN REF. 2).
I -> V (IN REF. 2).
I -> V (IN REF. 2).
W. 1497BlAB3F5FF2A4 CRC64;
                                                                                                                                                                                                                                                                                                                                                                CHLOROPLAST (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 55;
                                                                                                                                                                                                                                                                                                                                                                                    PROTEASE DO-LIKE 1. SERINE PROTEASE.
                                                                                                    -i- INDUCTION: By heat shock.
-i- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S2C.
-i- SIMILARITY: Contains 1 PDZ/DHR domain.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Conservative 17; Mismatches
                                                                                                                                                                                                                                                                   Interpro: IPR001478; PD2.
Interpro: IPR001440; Protease2C.
Interpro: IPR001054; Sr._protease_Iry.
Pfam. PF00595; PD2; 1.
Pfam. PF00689; trypsin; 1.
PRINTS; PR00884; PROTEASES2C.
PRINTS; SW00228; PD2; 1.
                                                                                                                                                                                                                                      EMBL; AP000371; BAB02539.1; -. EMBL; AP001302; BAB02539.1; JOINED.
                                                                                                                                                                                                                             EMBL; AF028842; AAC39436.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       46213 MW;
                                                                                                                                                                                                                                                                                                                                                       Hydrolase; Serine protease;
                                                                        O-PHENANTHROLINE.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Query Match
Best Local Similarity
Local 45; Conserva
                                                                                                                                                                                                                                                                                                                                                                                                                                                                     36
54
60
64
68
355
381
416
437 AA;
                                                                                                                                                                                                                                                                                                                                            PROSITE; PS50106;
                                                                                                                                                                                                                                                            MEROPS; S01.279;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CONFLICT
CONFLICT
CONFLICT
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                  ACT_SITE
ACT_SITE
ACT_SITE
CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CONFLICT
CONFLICT
CONFLICT
CONFLICT
                                                                                                                                                                                                                                                                                                                                                                          TRANSIT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                              DOMAIN
                                                                                                                                                                                                                                                                                                                                                                 RANSIT
ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ç
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           셤
```

RESULT 12

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL cutstation the European Bioinformatics Institute. There are the creative of its institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               68 HIVYHGAGTRIIASPKGPVTQMYTNVDKDLVGWQAPQGSRSL----TPCTCGSSDLYLV 122
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             123 TRHADVIPVRRR---GDSRGS-----LLSPRPISYLKGSSG--GPLLCPA------ 162
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Cancer Res. 55:2140-2149(1995).

-!- FUNCTION: MAY HAVE A FUNCTION IN MIGRATING CELLS.

-!- TISSUE SPECIFICITY: EXPRESSED IN BLOOD VESSELS. STRONGLY EXPRESSED

-- TISSUE CELLS, CYTOTROPHOBLASTS, AND POORLY DIFFERENTIATED COLON ADENOCARCINOMA CELLS POUND IN LYMPHATICS.

-!- SIMILARITY: Contains 8 WD repeats.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Gallardo J.A., Liotta L.A.; "Identification of a new immunoglobulin superfamily protein expressed in blood vessels with a heparin-binding consensus sequence.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Gaps
                                                                                                                                                                            Homo sapiens (Human).
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      -----GHAVGIFR----AAVSTRGVAKAVDFIPVESL 190
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        8.2%; Score 83; DB 1; Length 452; 25.3%; Pred. No. 2.8;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Conservative 13; Mismatches 47; Indels
                                                                                                                                                                                                                                                                                                                                                                                              Beckner M.E., Krutzsch H.C., Stracke M.L., Williams S.T.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                HEPARIN-BINDING (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     DA1413D25EB236C0 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               MIN. 603488...
GO: GO:0008201; F:heparin binding activity; TAS.
InterPro; IPR001680; WD40.
Fina. PF00400; WD40; 8.
SMART; SM00320; WD40; 8.
PROSITE; PS00678; WD_REPEATS_1; 1.
PROSITE; PS50082; WD_REPEATS_2; 6.
RROSITE; PS50082; WD_REPEATS_2; 6.
Repeat; WD repeat.
                                             15-JUL-1998 (Rel. 36, Created)
15-JUL-1998 (Rel. 36, Last sequence update)
28-FEB-2003 (Rel. 41, Last annotation update)
Anglo-associated migratory cell protein.
452 AA
                                                                                                                                                                                                                                                                                                                    SEQUENCE FROM N.A., AND PARTIAL SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          POLY-GLU.
WD 1.
WD 2.
WD 3.
WD 4.
WD 5.
WD 5.
WD 6.
                                                                                                                                                                                                                                                                                                                                              TISSUE-Brain;
MEDLINE-95262124; PubMed-7743515;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     49015 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         EMBL; M95627; AAA68889.1; -.
STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   18
77
138
180
220
220
261
306
404
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  PIR; I39383; I39383.
Genew; HGNC:18; AAMP.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Local Similarity
ses 42; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        452 AA;
                                                                                                                                                                                                                                                            NCBI_TaxID=9606;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bloinformatics and the EMBL outstation. the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       MEDIINE-92263794; PubMed-1316684;

Lacono-Connors L.C., Schmaljohn C.S.;

Lacono-Connors L.C., Schmaljohn C.S.;

Cloning and sequence analysis of the genes encoding the nonstructural proteins of Langat virus and comparative analysis with other flaviviruses.

1. CATALYTIC ACTIVITY: Hydrolysis of four peptide bonds in the viral precursor polyprotein. commonly with Asp or Glu in the P6 position, Cys or Thr in P1 and Ser or Ala in P1'.

1. CATALYTIC ACTIVITY: N nucleoside triphosphate - N diphosphate +
                                                                                                                        01-APR 1993 (Rel. 25, Created)
01-FEB-1994 (Rel. 28, Last sequence update)
01-FEB-1994 (Rel. 28, Last sequence update)
01-FEB-1003 (Rel. 41, Last annotation update)
Genome polyprotein [Contains: Capsid protein C (Core protein);
ENVelope protein Rem: Marrix protein (Envelope protein M; Major
ENVelope protein E: Nonstructural protein NS1; Nonstructural protein
NS2A; Nonstructural protein NS2B; Helicase/protease (EC 3.4.21.98)
(NS3); Nonstructural protein NS4B; Nonstructural protein
directed RNA polymerase (EC 2.7.7.48) (NS5)].
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A LIPOPROTEIN BYNTELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS: PROPEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MRNA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                    "Sequence of the genes encoding the structural proteins of the low-virulence tick-borne flaviviruses Langat TP21 and Yelantsev."; Virology 185:891-895(1991).
                                                                                                                                                                                                                                                                                         Langat virus (strain TP21).
Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                  [1]
SEQUENCE OF 1-776 FROM N.A.
MEDLINE-92074260; PubMed-1720591;
Mandl C.W., Iacono-Connors L., Wallner G., Holzmann H., Kunz C.,
Helnz F.X.;
296 GSVDCQAKLVSATTGKVVGVFRPETVASQPSLGEGEESESNSVESL 341
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Flavi_capsidc.
Flavi_glycoprotE.
Flavi_helicase.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            EMBL; M73835; AAA02740.1; ALT_TERM.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   IPRO00069; Flavi, M.
IPR001157; Flavi, NS2A.
IPR000752; Flavi, NS2B.
IPR000404; Flavi, NS2B.
IPR000404; Flavi, NS4A.
IPR0001228; Flavi, NS4B.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 IPR002535; Flavi_propep
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    FtsJ.
Hellcase_C
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 EMBL; S35365; AAB22165.1;
PIR; A42545; A42545.
                                                                                           STANDARD:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Interpro; IPR001410;
Interpro; IPR001122;
Interpro; IPR000336;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   IPR001850;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      IPR001650;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         HSSP; P14336; 15VB
                                                                                                                                                                                                                                                                                                                                               NCBI_TaxID=31638;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         MEROPS; S07.001;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  (RNA)(N).
                                                   RESULT 13
POLG_LANVT
ID POLG_LANVT
AC P29837;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  InterPro;
InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                InterPro;
                                                                                                               a
```

```
1496 GCSEGRSDSRPLDVKNGVYRIYTPGLLMGQRQIGVGYGAKGVLHTMWHVTRGAALLVDGV 1555
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1556 AVGP---YWADVREDVVCYGGAWSLESRWRGETVQVHAFPPG-RAHETHQCQPGELIL-- 1609
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               82 PKGPVTQMYINVDKDLV-----GWQA-----PQGSRSLTPCTCGSSDLYLVT 123
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   30 GCQETSOTGRDKNQVEGEVQIVSTA----TQTFLATSINGVLWTVYH---GACTRTIAS 81
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      PARALIS STANDARD DELLOCATOR DELLOCATION DE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              POTENTIAL.
POTENTIAL.
POTENTIAL.
POTENTIAL.
POTENTIAL.
POTENTIAL.
BY SIMILARITY.

                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             -----ENGREMGAI----PIDLAKGTSGSPIMNSQGEVVGLYGNGLKT 1648
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                124 RHADVIPVRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVST 175
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       DB 1; Length 3414;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ATP (POTENTIAL).
DEAH BOX.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  7.8%; Score 79.5; D
22.7%; Pred. No. 69;
ive 21; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          PD001556; Flavi_glycoprotE; 1. PD001496; Flavi_NS1; 1.
Interpro; IPR007095; RNA_pol_DS_PS.
InterPro; IPR007094; RNA_pol_PSvir.
Pfam; PP01003; Flavi_capsid; 1.
Pfam; PF02832; Flavi_glycop_c; 1.
Pfam; PF03869; Flavi_glycoprot; 1.
Pfam; PF00869; Flavi_halicase; 1.
                                                                                                                                                                                                                                                                                       Pfam; PF01004; Plavi, M; 1. Pfam; PF00948; Plavi, MS1; 1. Pfam; PF01005; Plavi, MS24; 1. Pfam; PF01005; Plavi, MS28; 1. Pfam; PF01349; Plavi, MS48; 1. Pfam; PF00372; Plavi, MS5; 1. Pfam; PF00572; Plavi, MS5; 1. Pfam; PF01570; Plavi, MS5; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Pfam; PF01728; FtsJ; 1.
Pfam; PF00271; helicase_C; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            378017
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SMART; SM00490; HELICC;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SMART; SM00487; DEXDc;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            3414 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Local Similarity
ses 39; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1688
1779
103
262
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            POLG_TBEVW
ID POLG_TBEVW
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              TRANSMEM
TRANSMEM
DISULFID
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         DISULFID
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             NP_BIND
SITE
TRANSMEM
TRANSMEM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             DISULFID
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            DISULFID
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     DISULFID
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ProDom;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ProDom;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Best Loc
Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              RESULT 14
          SOLUTION NEW TOUCH STREET STRE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   9
```

æ

STANDARD;

```
Flavi_glycoprotE.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               DISULFID
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           DISULFID
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         DOMAIN
NP_BIND
SITE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     DISCLFID
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       DISULFID
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            FRANSMEM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     PRANSMEM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               FRANSMEM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           RANSMEM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CHAIN
PROPEP
CHAIN
CHAIN
CHAIN
CHAIN
CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMEL outstation the European Bioinformatics Institute. There are no restrictions on its was by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/
                                                                                                                                                                                                                                                                                                                                                                                                       STRAIN-Neudoerfl;
MEDLINE-96036491; PubMed-7483260;
Walliner G., Handl C.W., Kunz C., Heinz F.X.;
"The flavivirus 3' noncoding region: extensive size heterogeneity independent of evolutionary relationships among strains of tick-borne
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    01-JAN-1990 (Rel. 13, Created)
30-MAY-2000 (Rel. 39, Last sequence update)
30-MAY-2000 (Rel. 39, Last sequence update)
30-MAY-2000 (Rel. 4), Last annotation update)
Genome polyprotein [Contains: Capsid protein C (Core protein); Matrix
protein (Envelope protein M); Major envelope protein E; Nonstructural
proteins NSI, NS2B, NS2B, NS4B, and NS4B; Protease/helicase
(EC 3.4.21.98) (NS3); RNA-directed RNA polymerase (EC 2.7.7.48)
(NS5)].
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                STRAIN-Neudoerfl:
MEDLINE-88322870; PubMed-3413985;
Mandl C.W., Heinz F.X., Kunz C.;
"Sequence of the structural proteins of tick-borne encephalitis virus (western subtype) and comparative analysis with other flaviviruses.";
Virology 166:197-205(1988).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS) OF 281-680.
MEDLINE-95272700; PubMed-7753193;
Rey F.A., Heinz F.X., Mandl C.W., Kunz C., Harrison S.C.;
"The envelope glycoprotein from tick-borne encephalitis virus at 2-A
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A LIPOPROTEIN BYNELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS: PROFIEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MRNA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               STRAIN-Neudoerfl;
MEDLINE-90051080; PubMed-2554575;
Mandl C.W., Heinz F.X., Stoeckl E., Kunz C.;
"Genome sequence of tick-borne encephalitis virus (Western subtype)
and comparative analysis of nonstructural proteins with other
flaviviruses.";
                                                                                                                                                                                                                                     Tick-borne encephalitis virus (Western subtype) (TBEV).
Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   PDB; 1SVB; 10-JUN-96.
MEROPS; SO7.UPW; --
InterPro; IPRO01410; DEAD.
InterPro; IPRO01122; Flavi_capsidC.
                                                                                                                                                                                                                                                                                                                                                                              SEQUENCE FROM N.A., AND REVISIONS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            SEQUENCE OF 767-3414 FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   EMBL; U27495; AAA86870.1; -. PIR; A31052; GNWVNE.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Virology 173:291-301(1989).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            SEQUENCE OF 1-779 FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                encephalitis virus.";
Virology 213:169-178(1995).
                                                                                                                                                                                                                                                                                                                           WCBI_TaxID=11088;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               [RNA](N)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   resolution."
                                                                                                                                                                                                                                                                                                 Flavivirus.
```

```
Probom; PEO0155 Flavi_1/2/1/20protE; 1.
Probom; PEO01496; Flavi_NS1; 1.
SNART; SN00497; DEXDC: 1.
SNART; SN00490; HELIC: 1.
Polyprotein; Glycoprotein; Transferase; RNA-directed RNA polymerase; Core protein; Coat protein; Envelope protein; Hydrolase; Helicase; ATP-binding; Transmembrane; Nonstructural protein; 3D-structure.
INII_MET 1 REMOVED FROM CAPSID PROTEIN C BY THE CELLULAR AMINOPEPTIDASE.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                HYDROPHOBIC SIGNAL SEQUENCE (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              RNA-DIRECTED RNA POLYMERASE (NS5).
INVOLVED IN FUSION.
ATP (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            MAJOR ENVELOPE PROTEIN E.
NONSTRUCTURAL PROTEIN NS.1.
NONSTRUCTURAL PROTEIN NS.2.
NONSTRUCTURAL PROTEIN NS.2.
PROTEASE/RELICASE (NS.3).
NONSTRUCTURAL PROTEIN NS.4.
NONSTRUCTURAL PROTEIN NS.4.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ENVELOPE GLYCOPROTEIN M.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (GLCNAC.
(GLCNAC.
(GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CAPSID PROTEIN C.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    POTENTIAL.
POTENTIAL.
POTENTIAL.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
               Interpro; IPR001850; Flavi_helicase.
Interpro; IPR000069; Flavi_M.
Interpro; IPR001057; Flavi_NSI.
Interpro; IPR00157; Flavi_NSI.
Interpro; IPR000487; Flavi_NS2A.
Interpro; IPR000487; Flavi_NS4A.
Interpro; IPR001528; Flavi_NS4B.
Interpro; IPR001528; Flavi_NS4.
Interpro; IPR001528; Flavi_NS6.
Interpro; IPR001535; Flavi_propep.
Interpro; IPR001650; Helicase_C.
Interpro; IPR001650; Helicase_C.
Interpro; IPR001697; FNA_pol_DS_PS.
Interpro; IPR001094; FNA_pol_DS_PS.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              DEAH BOX
                                                                                                                                                                                                                                                                                      Pfam; PF01003; Flavi_capsid; 1. Pfam; PF02832; Flavi_capsid; 1. Pfam; PF02832; Flavi_calycoprot; 1. Pfam; PF00949; Flavi_biloase; 1. Pfam; PF01004; Flavi_Ms; 1. Pfam; PF01005; Flavi_Ms; 1. Pfam; PF01005; Flavi_Ms2A; 1. Pfam; PF01350; Flavi_Ms2A; 1. Pfam; PF01350; Flavi_Ms2A; 1. Pfam; PF01350; Flavi_Ms4B; 1. Pfam; PF01349; Flavi_Ms4B; 1. Pfam; PF01728; Ftsy; 1. Pfam; PF01728; Ftsy; 1. Pfam; PF01728; Ftsy; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              7.1111328
13.22210
13.22210
13.22210
14.141
17.822
17.822
17.822
17.822
17.822
17.822
17.822
17.822
17.822
17.822
17.822
17.822
17.822
17.822
17.822
17.822
17.822
17.822
17.822
17.822
17.822
17.822
17.822
17.822
17.822
17.822
17.822
17.822
17.822
17.822
17.822
17.822
17.822
17.822
17.822
17.822
17.822
17.822
17.822
17.822
17.822
17.822
17.822
17.822
17.822
17.822
17.822
17.822
17.822
17.822
17.822
17.822
17.822
17.822
17.822
17.822
17.822
17.822
17.822
17.822
17.822
17.822
17.822
17.822
17.822
17.822
17.822
17.822
17.822
17.822
17.822
17.822
17.822
17.822
17.822
17.822
17.822
17.822
17.822
17.822
17.822
17.822
17.822
17.822
17.822
17.822
17.822
17.822
17.822
17.822
17.822
17.822
17.822
17.822
17.822
17.822
17.822
17.822
17.822
17.822
17.822
17.822
17.822
17.822
17.822
17.822
17.822
17.822
17.822
17.822
17.822
17.822
17.822
17.822
17.822
17.822
17.822
17.822
17.822
17.822
17.822
17.822
17.822
17.822
17.822
17.822
17.822
17.822
17.822
17.822
17.822
17.822
17.822
17.822
17.822
17.822
17.822
17.822
17.822
17.822
17.822
17.822
17.822
17.822
17.822
17.822
17.822
17.822
17.822
17.822
17.822
17.822
17.822
17.822
17.822
17.822
17.822
17.822
17.822
17.822
17.822
17.822
17.822
17.822
17.822
17.822
17.822
17.822
17.822
17.822
17.822
17.822
17.822
17.822
17.822
17.822
17.822
17.822
17.822
17.822
17.822
17.822
17.822
17.822
17.822
17.822
17.822
17.822
17.822
17.822
17.822
17.822
17.822
17.822
17.822
17.822
17.822
17.822
17.822
17.822
17.822
17.822
17.822
17.822
17.822
17.822
17.822
17.822
17.822
17.822
17.822
17.822
17.822
17.822
17.822
17.822
17.822
17.822
17.822
17.822
17.822
17.822
17.822
17.822
17.822
17.822
17.822
17.822
17.822
17.822
17.822
17.822
17.822
17.822
17.822
17.822
17.822
17.822
17.822
17.822
17.822
17.822
17.822
17.822
17.822
17.822
17.822
17.822
17.822
17.822
17.822
17.822
17.822
17.822
17.822
17.822
17.822
17.822
17.822
17.822
17.822
17.822
17.822
17.822
17.822
17.822
17.822
17.822
17.822
17.822
17.822
17.822
17.822
17.822
17.822
17.822
17.822
17.822
17.822
17.822
17.822
17.822
17.822
17.822
17.822
17.822
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              HELIX
STRAND
STRAND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   STRAND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 TURN
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@lsb-sib.ch).
                                                                                                                                                                                                                                                                                   TISSUE-EXTRACEMENT OF THE SAME M.H., Baker C.B.; MCDOWELL K.J., Adams W.H., Baker C.B.; MCDOWELL K.J., AMARIDAS DE SHARL/GEBBAK/DDBJ databases.

-!- FUNCTION: TRANSFERRINS ARE INON BINDING TRANSPORT PROTEINS WHICH CAN BIND TWO ATOMS OF FERRIC IRON IN ASSOCIATION WITH THE BINDING OF AN BAILON, USCARABE FOR THE TRANSPORT OF IRON FROM SITES OF ASSORPTION AND HEND DECRADATION TO THOSE OF STORAGE AND UTILIZATION. SERMY TRANSFERRIN MAY ALSO HAVE A FURTHER ROLE IN STIMULATING CELL PROLIFERATION.
                  01-AUG-1992 (Rel. 23, Created)
01-AUG-1992 (Rel. 23, Last sequence update)
28-FEB-2003 (Rel. 41, Last annotation update)
Serotransferrin precursor (Transferrin) (Siderophilin) (Beta-1-metal
binding globulin).
                                                                                                                  Equus caballus (Horse).
Bukaryota, Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Euteria; Perissodactyla; Equidae; Equus.
NCBI_TaxID=9796;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Iransport; Iron transport; Glycoprotein; Metal-binding; Repeat;
                                                                                                                                                                                                                                                                                                                                                                                                                  -1- SUBCNIT: Monomer.
-1- SUBCRILIAR LOCATION: Secreted.
-1- TISSUE SPECIFICITY: Expressed in liver: secreted in plasma.
-1- DOMAIN: COMPOSED OF TWO HOMOLOGOUS DOMAINS.
-1- SIMILARITY: BELONGS TO THE TRANSFERRIN FAMILY.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        BY SIMILARITY.
SEROTRANSFERRIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SIMILARITY.
Y SIMILARITY.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            SIMILARITY.
SIMILARITY.
SIMILARITY.
SIMILARITY.
SIMILARITY.
      706 A.A.
                                                                                                                                                                                      SEQUENCE FROM N.A.
MEDILNE-9377958; PubMed-8504171;
Carpenter M.A., Broad T.E.;
"The cDNA sequence of horse transferrin.";
Biochim. Biophys. Acta 1173:230-232(1993).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SIMILARITY.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SIMILARITY.
SIMILARITY.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Interpro; IPR001156; Transferrin. Pfam; PF00402; TRANSFERIN; 2. PRINTS; PR00422; TRANSFERIN. SMART; SM00094; TR_FER; 2. PROSITE; PS00205; TRANSFERRIN. 1; 2. PROSITE; PS00205; TRANSFERRIN. 1; 2. PROSITE; PS00206; TRANSFERRIN. 2; 2. PROSITE; PS00206; TRANSFERRIN. 3; 2.
      PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      EMBL; M69020; AAA30958.1; -. EMBL; U21127; AAA63684.1; -. PIR; S33761; S33761. HSSP; P02787; IABE.
      STANDARD:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         TRFE_HORSE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              DISULFID
DISULFID
DISULFID
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      DISULFID
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 DISCLETD
DISCLETD
DISCLETD
DISCLETD
DISCLETD
DISCLETD
DISCLETD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          DISULFID
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Signal.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SIGNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   REPEAT
REPEAT
      81 SPKGPVTQMYTNVDKDLV------GWQA-----PQGSRSLTPCTCGSSDLYLV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1555 AVAGP---YWADVREDVYCYGGAWSLEEKWKGETVQVHAFPPG-RAHEVHQCQPGELILD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 34 TSQTGRDKNQ-----VEGEVQIVSTA-----TQTFLATSINGVLWTVYH---GAGTRTIA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 T------GRKLGAIPIDLVKGTSGSPILNAGGVVGLYGNGKT 1648
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 123 TRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVST 175
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      50;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Score 79; DB 1; Length 3414; Pred. No. 77;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      eo; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Query Match 7.8%; Score 79; DB Best Local Similarity 23.7%; Pred. No. 77; Matches 41; Conservative 22; Mismatches
   1611
STRAND
TOTRA
STRAND
TOTRA
TOTR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                TURN
STRAND
STRAND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      STRAND
```

RESULT 15 TRFE\_HORSE

a

ð

g

õ

셤

à

```
T DISULFID 498 692 BY SIMILARITY.

T DISULFID 508 522 BY SIMILARITY.

T DISULFID 590 604 BY SIMILARITY.

T DISULFID 590 604 BY SIMILARITY.

T DISULFID 590 604 BY SIMILARITY.

T METAL 79 79 IRON 1 (BY SIMILARITY).

T METAL 209 209 IRON 1 (BY SIMILARITY).

T METAL 209 209 IRON 1 (BY SIMILARITY).

T METAL 413 413 413 HON 2 (BY SIMILARITY).

T METAL 449 449 1RON 2 (BY SIMILARITY).

T METAL 544 544 IRON 2 (BY SIMILARITY).

T METAL 612 612 BY SIMILARITY).

T METAL 612 612 IRON 2 (BY SIMILARITY).

T BINDING 480 480 IRON 2 (BY SIMILARITY).

T RINDING 480 480 ANION (POTENTIAL).

T CARBOHYD 515 515 515 515 515 515 515 515 FOOL 785 500 FOOL 785
```

parch completed: August 30, 2003, 19:13:48
btime : 11.7567 secs

hepatitis chepatitis c

091rs1 091rq8 091rt1 091rt1 091rt2 091rs3 091rt4 091rs2 091rs2

hepatitis on hepatitis of hepat

Q91rt8 O70818 Q91rr9 Q9dte2 Q99au2

```
SEQUENCE FROM N.A.
MEDILINE-203344; PubMed-10864644;
MEDILINE-203344; PubMed-10864644;
Lai V.C., Zhong W., Skelton A., Ingravallo P., Vassilev V.,
Donis R.O., Hong Z., Lau J.Y.;
"Generation and characterization of a hepatitis C virus NS3 protease-dependent bovine viral diarrhea virus.";
J. Virol. 74:6339-6347(2000).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
Pestivirus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   the EMBL/GenBank/DDBJ databases
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Created)
Last sequence update)
Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                          PRT; 4040 AA
                                                                                                                                                                                                                                                                                                                                                                                                          ALIGNMENTS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               InterPro; IPR000280; CDvir_endptseP80.
InterPro; IPR000140; MCV_NS3.
InterPro; IPR001410; MCV_NS3.
InterPro; IPR001410; MCV_NS3.
InterPro; IPR001505; Mpl.DNR_binding.
InterPro; IPR001005; Myb_DNR_binding.
InterPro; IPR001005; NNA_pol_DS_PS.
InterPro; IPR007095; RNA_pol_DS_PS.
InterPro; IPR007094; RNA_pol_DS_PS.
Fam; PF02907; HCV_NS3: 1.
Ffam; PF00271; helicase_C; 1.
Ffam; PF00271; helicase_C; 1.
091RS1
091RNB
091RNB
091RR0
031RR0
091RR3
091RS3
091RS4
091RS7
                                                                                                                                                                                                                                                                                                                            070818
Q91RR9
Q9DTE2
Q99AU2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Lai V.C.H., Hong Z.;
Submitted (MAY-2000) to the
EMBL; AF268278; AAF82566.1;
HSSP; P26663; 1JXP.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Q9IFH8;
01-OCT-2000 (TIEMBLIE). 15,
01-OCT-2000 (TIEMBLIE). 15,
01-MAR-2003 (TIEMBLIE). 23,
  PRELIMINARY;
 Genome polyprotein.
Mucosal disease virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        NCBI_TaxID-11099;
 MEROPS; S31.001;
                                                                                                                                                                                                                                                              838.5
838.5
838.5
838.5
837.5
837.5
837.5
841
840.5
840
839.5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Q9IFH8
                                                                                                                                                                                                                                                                                                                                                                                                                                               RESULT 1
Q9IFH8
 PAC DE REPRESENTATION DE LA PROPERTA DE LA PROPERTA DE LA PRESENTA DEL PRESENTA DE LA PRESENTA DE LA PRESENTA DEL PRESENTA DE LA PRESENTA DE 
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            091fhe mucosal dis
03679 hepatitis c
091f56 hepatitis c
091e5 hepatitis c
03608 hepatitis c
036608 hepatitis c
090xx5 hepatitis c
091xx8 hepatitis c
091xx6 hepatitis c
091xx7 hepatitis c
091xx7 hepatitis c
091xx7 hepatitis c
091xx8 hepatitis c
091xx8 hepatitis c
091xx8 hepatitis c
091xx8 hepatitis c
                                                                                                   (without alignments)
1352.314 Million cell updates/sec
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Pred. No. 1s the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.
                                                                                     August 30, 2003, 19:00:22; Search time 37.5921 Seconds
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Description
                                                                                                                                        US-09-965-594-18
1017
1 MKKKGSVVIVGRINLSGDTA.....VAKAVDFIPVESLETIMRSP
          GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd
                                                                                                                                                                                                                                                          Total number of hits satisfying chosen parameters:
                                                                                                                                                                                                                                   830525 seqs, 258052604 residues
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SUMMARIES
                                                                                                                                                                                                                                                                                                                          Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

    protein search, using sw model

                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              03463
036608
09PWX5
09PWU9
091RR8
091RT5
091RR5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Q91FH8
Q36579
Q81756
Q91FES
Q9ELS8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     091RT9
091RR3
091RR4
                                                                                                                                                                                              BLOSUM62
Gapop 10.0 , Gapext 0.5
                                                                                                                                                                                                                                                                                                                                                                                                                   sp_fungi:*
sp_human:*
sp_invertebrate:*
sp_mammal:*
sp_mhc:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   sp_vertebrate:*
sp_unclassified:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             sp_rvirus:*
sp_bacteriap:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 sp_organelle:*
sp_phage:*
                                                                                                                                                                                                                                                                                     Minimum DB seq length: 0
Maximum DB seq length: 2000000000
                                                                                                                                                                                                                                                                                                                                                                                           sp_archea:*
sp_bacterla:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           sp_plant: *
sp_rodent: *
sp_virus: *
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        sp_archeap:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             DB
                                                                                                                                                                                                                                                                                                                                                                               SPTREMBL_23:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Query
Match Length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          2436
3011
3011
3011
3011
3015
3015
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    181
181
181
181
181
181
181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Scoring table:
                                                                                                                                                      Perfect score:
Sequence:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               872.5
858.5
854.5
854.5
854.5
853.5
851.5
851.5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   849
847
847
847
847
846
846
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Score
                                                               OM protein
                                                                                                                                                                                                                                   Searched:
                                                                                                                                                                                                                                                                                                                                                                               Database
                                                                                        on:
```

~

```
PF01543;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       081756
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     RESULT 3
081756
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      DER KRENK OCCOSETATOR OF A PROPERTY OF A PRO
                   SA WEEN STANFASTER SO SENSOR S
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ద
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Ĥ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           62 SINGVLWTVYHGAGTRIIASPRGPVTQMYINVDKDLVGWQAPQGSRSLIPCTCGSSDLYL 121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      10 GSVVIVGRIVLSGSGSITACAQQTRGLLGCKITSLTGRDKNQVEGEVQIVSTATQTFLAT 69
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      5 GSVVIVGRINLSGD---TAYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVSTATQTFLAT 61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              S. E.
PRINTS; PR00729; CDVENDOPTASE.
SMART; SM00487; DEXDC; 1.
SRART; SM00490; HELICO; 1.
PROSITE; PS00037; MTB_1; 1.
PROSITE; PS05507; RDRP_POSITIVE; 1.
PROSITE; PS05507; RDRP_POSITIVE; 1.
PROSITE; PS05531; RNASE_T12, 1.
ATP-binding; Helicase; Hydrolase; Nonstructural protein; Polyprotein; BRA-directed RNA, Polymerase; Transferase.
SEQUENCE 4040 AA; 453073 MW; ADE87791D05589DC CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      "Transmission of hepatitis C by intrahepatic inoculation with transcribed RNA.";
Science 277:570-574(1997).
Science 277:570-574(1997).
I- SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A LIPOPROTEIN ENVELOPE CONSISTS OF TWO PROTEINS:
PROPEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MRNA (BY SIMILARITY).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               VITUSES; SSRNA positive-strand viruses, no DNA stage; Flaviviridae; Hepacivirus.
NCBL_TaxID-11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         MEDLINE-97373636; PubMed-9228008;
Kolykhalov A.A., Agapov E.V., Blight K.J., Mihalik K., Feinstone
Rice C.M.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Length 4040;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ά,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Score 872.5; DB 12; Length
Pred. No. 1e-75;
5; Mismatches 13; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Last sequence update)
Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 PRT; 3011 A.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Created)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          HCV_NS5a.
HCV_RdRP.
Helicase_C.
RNA_pol_DS_PS.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      HCV_capsid.
HCV_core.
HCV_env.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            HCV_NS1.
HCV_NS2.
HCV_NS3.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        HCV_NS4b.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   HCV_NS4a.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             VDFIPVESLETTMRS 196
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              01-JAN-1998 (TEMBLE). 05, 01-JAN-1998 (TEMBLE). 05, 01-MAR-2003 (TEMBLE). 23, Genome Polyprotein. Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               EMBL; AF009606; AAB66324.1;
HSSP; P27958; 1HEI.
InterPro; IPR001410; DEAD.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Query Match
Best Local Similarity 89.2%;
Matches 174; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              UDFIPVENLETTES 204
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          IPR001650; F
IPR007095; E
IPR007094; F
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 InterPro; IPR002522;
InterPro; IPR002521;
InterPro; IPR002519;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 IPR002531;
IPR002518;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        IPR000745;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        IPR002868;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             IPR004109;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               IPR002166;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   InterPro;
InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             182
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    190
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 036579
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           RESULT 2
036579
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                A PART OF A PART
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ద
              DR W DR W DR W B DR W B
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    a
```

```
54 ATQTFLATSINGVLWTVYHGAGTRTIASPKGPVTQMYTNVDKDLVGWQAPQGSRSLTPCT 113
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              3 KKGSVVIVGRIN-----LSGDTAYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVST 53
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        9; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                           protein; Glycoprotein; Helicase;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
Hepacivirus.
NCBI_TaxID-11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  DB 12; Length 3011;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Choo Q.-L., Richman K., Han J.;
"The nucleotide sequence of the Hepatitis C viral genome.";
Submitted (MAY-1990) to the EMBL/GenBank/DDBJ databases.
EMBL: M32084; AAA456777.1; -.
HSSP: P27958; 1A1V.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Ouery Match 84.4%; Score 858.5; DB 12; Length Best Local Similarity 82.8%; Pred. No. 1.6e-74; Matches 169; Conservative 10; Mismatches 16; Indels
                                                                                                                                                                                                                                                                                                                                               PROSITE; PS50507; RDRP_POSITIVE: 1.
PROSITE; PS50521; RDRP_VIRAL; 1.
ATP-binding; Coat protein; Envelope protein; Glycoproteil
Hydrolass; Nonstructural protein; Polyprotein;
RNA-directed RNA Polymerase; Transferase; Transmembrane.
SEQUENCE 3011 AA; 327182 MW; E2E0EE809C63C1B9 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    01-NOV-1996 (TIEMBLrel. 01, Created)
01-NOV-1996 (TIEMBLrel. 01, Last sequence update)
11-NAR-2003 (TIEMBLrel. 23, Last annotation update)
Genome polyprotealn (Fragment).
Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       PRT; 2436 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  InterPro; IPR002531; HCV_NS1.
InterPro; IPR002518; HCV_NS2.
InterPro; IPR004109; HCV_NS3.
InterPro; IPR00145; HCV_NS4.
InterPro; IPR001490; HCV_NS4b.
InterPro; IPR001868; HCV_NS5a.
InterPro; IPR001650; HCV_RSP.
InterPro; IPR001650; Helicase_C.
InterPro; IPR001650; Helicase_C.
InterPro; IPR007095; RNA_pol_DS_PS.
InterPro; IPR007095; RNA_pol_PSvIr.
Pfam; PF0150; HCV_NS1; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     174 STRGVAKAVDFIPVESLETTMRSP 197
Pfam; PF01542; HCV_core; I Pfam; PF01539; HCV_core; I Pfam; PF01542; HCV_core; I Pfam; PF01569; HCV_mS1; I Pfam; PF01506; HCV_mS3; I Pfam; PF01006; HCV_mS3; I Pfam; PF01001; HCV_mS4; I Pfam; PF01001; HCV_mS5a; I Pfam; PF01505; HCV_mS5a; I Pfam; PF005948; Viral_case_C; I Pfam; PF005948; Viral_case_C; I ProDom; PD186062; HCV_mS1; I SMART; SM0487; DEXDC; I.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        InterPro; IPR001410; DEAD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SEQUENCE FROM N.A.
```

ä

```
54 ATQTFLATSINGVLWTVYHGAGTRTIASPKGPVTQMYTNVDKDLVGWQAPQGSRSLTPCT 113
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    114 CGSSDLYLVTRHADVIPVRRKGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAV 173
                                                                                                                                                                                                                                                                                                                                                                                                              3 KKGSVVIVGRIN-----LSGDTAYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVST
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             9; Gaps
                                                                                                                                                                                                                                          PROSITE: PS00190: CYTCCHROME_C; 1.
PROSITE: PS00190: RDRP_POSITIVE; 1.
PROSITE: PS00501; RDRP_VIRAL; 1.
ATP-binding: Coat protein: Envelope protein; Glycoprotein; Helicase; Hydrolase; Monstructural protein; Polyprotein;
RNA-directed RNA polymerase; Transferase; Transmembrane.
SEQUENCE 3011 AA; 327124 MW; 2489CE74AC864E58 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
Hepacivirus
                                                                                                                                                                                                                                                                                                                                                         DB 12; Length 3011;
                                                                                                                                                                                                                                                                                                                                                         Ouery Match 84.0%; Score 854.5; DB 12; Length Best Local Similarity 82.8%; Pred. No. 3.9e-74; Matches 169; Conservative 9; Mismatches 17; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Last sequence update)
Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      PRT; 3011 A.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       174 STRGVAKAVDFIPVESLETTMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               01-WAR-2001 (TYEMBLREL. 16, Created)
01-WAR-2001 (TYEMBLREL. 16, Last sequ
01-WAR-2003 (TYEMBLREL. 23, Last and
Genome polyprotein.
           InterPro; IPR002166; HCV_RGRP.
InterPro; IPR001650; Helicase_C.
InterPro; IPR007095; RN_pol_DS_PS.
InterPro; IPR007094; RN_pol_PSvir.
Pfam; PF01543; HCV_capsid; 1.
Pfam; PF01543; HCV_core; 1.
Pfam; PF01559; HCV_env; 1.
Pfam; PF01569; HCV_env; 1.
Pfam; PF01589; HCV_env; 1.
                                                                                                                       Pfam; PF01538; HCV_NS2; 1. Pfam; PF012907; HCV_NS3; 1. Pfam; PF01006; HCV_NS4a; 1. Pfam; PF01506; HCV_NS5a; 1. Pfam; PF00271; helicase_C; 1. Pfam; PF00998; viral_RdRP; 1. ProDom; PD186062; HCV_NS1; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      PRELIMINARY;
                                                                                                                                                                                                                                    DEXDc;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              NCBI_TaxID-11103;
                                                                                                                                                                                                                                    SMART; SM00487;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    STRAIN-colonel;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Q9ELS8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           RESULT 5
Q9ELS8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     a
     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          à
                                                                                                                                                                                                                                                                                                       54 ATQTFLATSINGVLWTVYHGAGTRTIASPKGPVTQMYTNVDKDLVGWQAPQGSRSLTPCT 113
                                                                                                                                                                                                                                                                                                                                                          114 CGSSDLYLVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAV 173
                                                                                                                                                                                                                                                                                        3 KKGSVVIVGRIN------LSGDTAYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVST 53
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                SEQUENCE FROM N.A.
MEDLINE-21262212; Pubmed-11369872;
Lanford N.E., Lee H., Chavez D., Guerra B., Brasky K.M.;
Infectious cDNA clone of the hepatitis C virus genotype 1 prototype
                                                                                                                                                                                                                                                             9; Gaps
                                                                                                                                       Helicase;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Genome polyprotein.
Hepalitis C virus.
Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
Hepacivirus.
                                                                                                                                                                                                                                  84.0%; Score 854.5; DB 12; Length 2436; 82.8%; Pred. No. 2.9e-74; ative 9; Mismatches 17; Indels 9;
                                                                                                       PROSITE; PS50507; RDRP_POSITIVE; 1.
PROSITE; PS50521; RDRP_VIRAL; 1.
PROSITE; PS50521; RDRP_VIRAL; 1.
FYP-bInding; Coat protein; Envelope protein; Glycoprotein;
Hydrolase; Nonstructural protein; Polyprotein;
RNA directed RNA polymerase; Transferase; Transmembrane.
                                                                                                                                                                                           2436 2436
2436 AA: 264734 MW; D7B9872900BF3125 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Q9IFE5;
01-OCT-2000 (TrEMBLrel. 15, Created)
01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               PRT; 3011 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        HSSP; 72/303 LAAV.
INTERPTO; IPR000345; CYtC_heme_bind.
InterPro; IPR001410; DEAD.
InterPro; IPR001521; HCV_copsid.
InterPro; IPR002521; HCV_core.
InterPro; IPR002519; HCV_env.
InterPro; IPR002519; HCV_NS1.
InterPro; IPR002518; HCV_NS2.
InterPro; IPR004109; HCV_NS3.
InterPro; IPR004109; HCV_NS3.
InterPro; IPR001490; HCV_NS4.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                 735 CTRGVAKAVDFIPVENLETTMRSP 758
                                                                                                                                                                                                                                                                                                                                                                                                                                                   174 STRGVAKAVDFIPVESLETTMRSP 197
Pfam; PF02907; HCV_NS3; 1.
Pfam; PF01006; HCV_NS4s; 1.
Pfam; PF01001; HCV_NS4b; 1.
Pfam; PF01506; HCV_NS5s; 1.
Pfam; PF00271; hellcasc_C; 1.
Pfam; PF0098; Viral_RGRP; 1.
ProDom; P186062; HCV_NS1; 1.
SWART; SM00487; DEXDC; 1.
                                                                                                                                                                                                                                                   Best Local Similarity 82.8
Matches 169; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               PRELIMINARY;
                                                                                                                                                                                                                                                 Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      NCBI_TaxID-11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      seguence."
                                                                                                                                                                                             NON_TER
SEQUENCE
                                                                                                                                                                                                                                    Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Q9IFE5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 a
                                                                                                                                                                                                                                                                                                                                                                                                 ŏ
```

õ 셤 ò q

SON KEN DER KE

```
5'-terminal sequence of the hepatitis C virus genome."; J. Exp. Med. 60:167-177(1990).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           PROSITE; PS50507; RDRP_POSITIVE; 1.
PROSITE; PS50521; RDRP_VIRAL; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             InterPro; IPR002166; HCV_RdRP.
InterPro; IPR001650; Hellcase_C.
InterPro; IPR007095; RNA_pol_DS_PS.
InterPro; IPR007094; RNA_pol_PS_PS.
Pfam; PF01543; HCV_capsid; 1.
                                                                                                               STRAIN-HC-J1;
MEDLINE-92044440; PubMed-1658196;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     MEDLINE-94174722; PubMed-7510436;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            HCV_capsid.
HCV_core.
HCV_env.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Pfam; PF02907; HCV_NS3; 1.
Pfam; PF01006; HCV_NS4s; 1.
Pfam; PF01001; HCV_NS4b; 1.
Pfam; PF01506; HCV_NS5s; 1.
Pfam; PF00271; hclicase_C; 1.
Pfam; PF00998; viral_RdRP; 1.
Probom; P0186062; HCV_NS1; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     HCV_NS4A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        HCV_NS5a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                HCV_NS1.
HCV_NS2.
HCV_NS3.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              InterPro; IPR001410; DEAD.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  DEXDC;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 IPR004109; F
IPR000745; F
IPR001490; F
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               IPR002519;
IPR002531;
IPR002518;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               InterPro; IPR002522; InterPro; IPR002521;
                                                                                                                                                                                                                                                                                                                                                                    SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SEQUENCE FROM N.A.
                                                                                 SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                SEQUENCE FROM N.A
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SM00487;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     PF01560;
PF01538;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              PF01542;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          PF01539;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    STRAIN-HC-J1;
                                                                                                                                                                                                                                                                                                                                                                                                     STRAIN-HC-J1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            STRAIN-HC-J
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               InterPro;
InterPro;
InterPro;
InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Okamoto
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Pfam;
      ä
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      54 AIQTFLATSINGVLWTVYHGAGTRTIASPKGPVTQMYTNVDKDLVGWQAPQGSRSLTPCT 113
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              114 CGSSDLYLVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAV 173
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          53
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    3 KKGSVVIVGRIN-----LSGDTAYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVST
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              9; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                protein; Glycoprotein; Helicase;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Genome polyprotein.
Hepatitis C virus.
Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
Hepativirus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        DB 12; Length 3011;
                                                                                                                                                                                                                                                                                                                                                                                                                 Pfam; PFOLSS 9; MCV_env; 1.

R Pfam; PFOLSS 9; MCV_env; 1.

R Pfam; PFOLSS 8; MCV_env; 1.

R Pfam; PFOLSO 90; MCV_NS 2; 1.

R Pfam; PFOLSO 90; MCV_NS 3; 1.

R Pfam; PFOLSO 1: MCV_NS 4; 1.

R Pfam; PFOLSO 91; MCV_NS 9; 1.

R Pfam; PFOLSO 91; MCV_NS 9; 1.

R PFOLST 92; PSOLS 93; MCN_NS 9; 1.

R PROSTIE; PSOLS 93; MCN_NS 9; 1.

R PROSTIE; PSOLS 94; MCN_NS 9; 1.

R PROSTIE; PSOLS 95; MCN_NS 9; 1.

R PROSTIE 95; PSOLS 95; MCN_NS 9; 1.

R PROSTIES 95; MCN_
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 STRAIN-HC-J1;
MEDLINE-91013116; PubMed-2170712;
MAGMOCO H., Okada S., Suglyama Y., Yotsumoto S., Tanaka T.,
Yoshizawa H.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  84.0%; Score 854.5; DB 12; Length
82.4%; Pred. No. 3.9e-74;
.1ve 11; Mismatches 16; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Q03463 PRELIMINARY; PRT; 3011 AA. 003463; Ol.NOV-1996 (TrEMBLrel. 01, Created) 01-NOV-1996 (TrEMBLrel. 01, Last sequence update) 01-NAR-2003 (TrEMBLrel. 23, Last annotation update)
                                                                                                                                                                                              LACOT HOLDS A.

LECTOR INTO TO THE TO
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1185 CTRGVAKAVDFIPVENLETTMRSP 1208
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          174 STRGVAKAVDFIPVESLETTMRSP 197
                                                                                                                                                                      HCV_NS4a.
HCV_NS4b.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Best Local Similarity 82.4%
Matches 168; Conservative
                                                                                                                                       IPR004109;
IPR000745;
                        IPR002521;
                                                   FPR002519;
                                                                                                            IPR002518;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            NCBI_TaxID-11103;
                                                                                                            InterPro;
InterPro;
InterPro;
InterPro;
InterPro;
                                                   InterPro;
InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Query Match
```

ö a ò ద ò g ठ d RESULT 6

THE REAL REAL PROPERTY OF THE REAL PROPERTY OF THE

```
virus isolated
Okamoto H., Okada S., Sugiyama T., Kurai K., Iizuka H., Machida A., Miyakawa Y., Mayumi M.;
"Nucleotide sequences of the genomic RNA of hepatitis C virus isolated from a human cariters comparison with reported isolates for conserved and divergent regions.";
J. Gen. Virol. 72:2697-2704(1991).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 "Characterization and mapping of a B-cell immunogenic domain in hepatitis C virus E2 glycoprotein using a yeast peptide library."; virology 200:246-255(1994).
-1- SUBDMIT: THE VIRION OF THIS VIRUS IS A NUCLECCAPSID COVERED BY A LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS: PROTEIN M AND GLYCOPROTEIN E. THE NUCLECCAPSID IS A COMPLEX OF PROTEIN C AND MRNA (BY SIMILARITY).
EMBL, D10749; BAA01582.1; --
HSSP, P27958; 1HEI.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ATP-binding; Coat protein; Envelope protein; Glycoprotein; Helicase;
                                                                                                                                                                                                                                                                                                                                                                                                  WEDLINE-93117120; PubMed~1335573;
Okamoto H., Kanai N., Mishiro S.;
Frull-length nucleotide sequence of a Japanese hepatitis C virus
isolate (HC-1) with high homology to USA isolates.";
Nucleic Acids Res. 20:6410-6410{1992}.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Mink M., Benichou S., Madaule P., Tiollais P., Prince A., Inchauspe G.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Hydrolase; Nonstructural protein; Polyprotein;
RNA-directed RNA polymerase; Transferase; Transmembrane.
SEQUENCE 3011 AA; 327112 MM; 97E9052C0250463B CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Okamoto H.;
Submitted (DEC-1992) to the EMBL/GenBank/DDBJ databases.
```

```
PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    InterPro; IPR002521;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             IPR002522;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         InterPro; IPR002519;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       InterPro; IPR002531;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   IPR004109;
IPR000745;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          IPR001490;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           IPR002868;
IPR002166;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                IPR002518;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Genome polyprotein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              NCBI_TaxID-11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Hepacivirus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           InterPro;
InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               InterPro;
                                                                                                                                                                                               Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Q9PWX5;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            RESULT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 09PWX5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     DATE TO DESCRIPE TERMINATION OF THE PARTY OF
d
                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                        a
                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                          ä
                                                                                                                54 ATQTFLATSINGVLWTVYHGAGTRTIASPKGPVTQMYTNVDKDLVGWQAPQGSRSLTPCT 113
                                                                                                                                                                                                                                                   114 CGSSDLYLVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAV 173
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            MEDLINE-97385173; PubMed-9238047; Yanagal M.; Purcell R.H., Emerson S.U., Bukh J.; Yanagal M., Purcell R.H., Emerson S.U., Bukh J.; Transcripts from a single full-length cDNa clone of hepatitis C virus are infectious when directly transfected into the liver of a
                                                                                                                                                                                             ---- LSGDTAYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVST
                                                        9; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ∢
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Genome polyprotein.
Hepatitis C virus strain H77.
Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
Hepacivirus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Proc. Natl. Acad. Sci. U.S.A. 94:8738-8743(1997).

-i- SUBUNT: THE VIRION OF THIS VIRUS IS A NUCLECCAPSID COVERED BY LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS: PROTEIN M AND GLYCOPROTEIN B. THE NUCLECCAPSID IS A COMPLEX OF PROTEIN C AND MRNA (BY SIMILARITY).

EMBL; AR011751; AAB67036.1; -- HSSP; P27958; 1HEI.
                DB 12; Length 3011;
                                                        Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          01-JAN-1998 (TrEMBLrel. 05, Created)
01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
                                                        18;
              Score 853.5; DB 1
Pred. No. 4.9e-74;
8; Mismatches 18
                                                                                                                                                                                                                                                                                                                                                                                                                                                                     PRT; 3011 AA
                                                                                                                                                                                                                                                                                                                                                       Interpro; IPR001410; DEAD.
Interpro; IPR001522; HCV_capsid.
Interpro; IPR002521; HCV_capsid.
Interpro; IPR002521; HCV_core.
Interpro; IPR002519; HCV_env.
Interpro; IPR002519; HCV_NS2.
Interpro; IPR002519; HCV_NS2.
Interpro; IPR001409; HCV_NS3.
Interpro; IPR001409; HCV_NS4b.
Interpro; IPR001409; HCV_NS4b.
Interpro; IPR001569; HCV_NS4b.
Interpro; IPR001569; HCV_NS5a.
Interpro; IPR001569; Helicase_C.
Interpro; IPR001509; RNA_pol_DS_PS.
Interpro; IPR0017095; RNA_pol_DS_PS.
Interpro; IPR007095; RNA_pol_DS_PS.
Interpro; IPR007094; RNA_pol_DS_PS.
Interpro; IPR007094; RNA_pol_DS_PS.
                                                                                                                                                                                                                                                                                                                                  174 STRGVAKAVDFIPVESLETTMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              helicase_C; 1
                83.9%;
82.8%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     HCV_NS1; 1.

HCV_NS2; 1.

HCV_NS3; 1.

HCV_NS4a; 1.

HCV_NS4b; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Pfam; PF01542; HCV_core; 1.
Pfam; PF01539; HCV_env; 1.
Pfam; PF01560; HCV_NS1; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           HCV_NS5a; 1
                                                                                               KKGSVVIVGRIN----
                                     cal Similarity 82.8
169; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                       PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        [1]
SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   NCBI_TaxID=63746;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           PF01506;
PF00271;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            PF01538;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             chimpanzee.";
                Query Match
Best Local
                                                                                                                                                                                                                                                                                                                                                                                                                                                                       036608
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            036608
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Pfam;
Pfam;
Pfam;
Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Pfam;
                                                        Matches
                                                                                                                                                                                                                                                                                                                                                                                                                               RESULT 7
036608
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ö
                                                                                                                                  셤
                                                                                                                                                                            õ
                                                                                                                                                                                                               8
                                                                                                                                                                                                                                                       ö
                                                                                                                                                                                                                                                                                          ద
                                                                                                                                                                                                                                                                                                                                  õ
                                                                                                                                                                                                                                                                                                                                                                     셤
```

```
1125 GGSSDLYLVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGLFRAAV 1184
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  113
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       114 CGSSDLYLVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAV 173
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  54 ATQTFLATSINGVLWTVYHGAGTRTIASPKGPVTQMYTNVDKDLVGWQAPQGSRSLTPCT
                                                                                                                                                                                                                                                                                                                                                                                        3 KRGSVVIVGRIN------LSGDTAYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVST
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   BY A
Pfam; PF00998; Viral_RdRP; 1.
ProDom; PD186062; HCV_NS1; 1.
SWART; SMO0487; DEXDC: 1.
PROSITE; PS50501; RDRP_POSITIVE; 1.
PROSITE; PS50521; RDRP_VIRAL; 1.
ATP-binding; Coat protein; Envelope protein; Glycoprotein; Helicase; Hydrolase; Nonstructural protein; Polyprotein; Clycoprotein; New Adirected RNA polymerase; Transferase; Transmembrane.
SEQUENCE 3011 AA; 327112 MW; 0875E6B81CB5C198 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Repatitis C virus.
Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          BURN J.;
SUDMITTED (AUG-1999) to the EMBL/GenBank/DDBJ databases.
-!- SUBDNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY
LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
PROTEIN M AND GIYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF
PROTEIN C AND MRNA (BY SIMILARITY).
                                                                                                                                                                                                                                                                                DB 12; Length 3011;
                                                                                                                                                                                                                                                                                                                                    6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  MEDLINE-99420396; PubMed-10489358;
Yanaqi M., Purcell R.H., Emerson S.U., Bukh J.;
"Hepatitis C virus: an infectious molecular clone of a second squotype (2a) and lack of viability of intertypic la and 2a chimeras.";
                                                                                                                                                                                                                                                                           83.7%; Score 851.5; DB 12; Length
82.4%; Pred. No. 7.7e-74;
Live 10; Mismatches 17; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                01-MAY-2000 (TrEMBLrel. 13, Created)
01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               PRT; 3015 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 174 STRGVAKAVDFIPVESLETTMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     HCV_capsid.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            HCV_env.
HCV_NSI.
HCV_NS3.
HCV_NS4.
HCV_NS4b.
HCV_NS5a.
HCV_NS5b.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        EMBL; AF177040; AAF01182.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Virology 262:250-263(1999).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         HSSP; P27958; IHEI.
InterPro; IPR001410; DEAD.
                                                                                                                                                                                                                                                                                                          Local Similarity 82.49 ses 168; Conservative
```

(BY SIMILARITY).

```
Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Q91RR8
Q91RR8;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   RESULT 10
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Q91RRB
      q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            $88855558885
                                                                                                                                                                                                                                                                                                                                                                       54 ATQIFLATSINGVLWIVYHGAGTRIIASPKGPVIQMYINVDKDLVGWQAPQGSRSLIPCT 113
                                                                                                                                                                                                                                                                                                                                                                                                                                                       114 CGSSDLYLVTRHADVIPVRRRGDSRGSLLSPRPISYLRGSSGGPLLCPAGHAVGIFRAAV 173
                                                                                                                                                                                                                                                                                                                                                                                                                                1069 ATQTFLATCINGVCWTVYHGAGTRTIASPKGPVIQMYTNVDQDLVGWPAPQGSRSLTPCT
                                                                                                                                                                                                                                                                                                                                                          3 KKGSVVIVGRIN-----LSGDTAYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVST
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Submitted (AUG-1999) to the EMBL/GenBank/DDBJ databases.
-1- SUBDANT: THE VIRIN OF THIS VIRIN IS A NOCLEOCARSID COVERED BY A LIPOPROTEIN ENVELOPE. THE EMPELOPE. THE EMPELOPE THE PROPELOPE PROPERS.
PROTEIN H AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF
                                                                                                                                                                                              PROSITE: NO0487; DEXDC: 1.
PROSITE: PS00392; DEXDC: 1.
PROSITE: PS00392; DEXDC: 1.
PROSITE: PS0507; RDRP_POSITIVE: 1.
ATP-binding: Coat protein; Envelope protein; Glycoprotein; Helicase; Hydrolase; Nonstructural protein; Polyprotein;
RNA-directed RNA polymerase; Transferase; Transmembrane.
SEQUENCE: 3015 AA: 328159 MW; B7D23BC1F190663A CRC64;
                                                                                                                                                                                                                                                                                                                                  9; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   MEDLINE-99420396; PubMed-10489358; Yanaqi M., Purcell R.H., Emerson S.U., Bukh J.; Yanaqi M., Purcell R.H., Emerson S.U., Bukh J.; Hepatitia C virus: an infectious molecular clone of a second major genotype (2a) and lack of viability of intertypic la and 2a chimeras.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae; Hepacivirus.
NCBL_TaxID=11103;
                                                                                                                                                                                                                                                                                                            DB 12; Length 3015;
                                                                                                                                                                                                                                                                                                          83.7%; Score 851.5; DB 12; Length
82.4%; Pred. No. 7.7e-74;
Live 10; Mismatches 17; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Last sequence update)
Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         PRT; 3015 A.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     InterPro: IPR001650; Helicase_C.
InterPro: IPR00129; Pyridoxal_deC.
InterPro: IPR007095; RNA_pol_DS_PS.
InterPro: IPR007095; RNA_pol_DS_PS.
InterPro: IPR007094; RNA_pol_DS_PS.
InterPro: IPR007094; RNA_pol_PSvir.
Pfam; PF01543; HCV_capsid; I.
Pfam; PF01542; HCV_capsid; I.
Pfam; PF01560; HCV_NS1; I.
Pfam; PF01560; HCV_NS1; I.
Pfam; PF01000; HCV_NS3; I.
Pfam; PF01000; HCV_NS4; I.
Pfam; PF00100; HCV_NS4; I.
Pfam; PF00271; helicase_C; I.
Pfam; PF00271; helicase_C; I.
Pfam; PF00271; helicase_C; I.
Pfam; PF00271; helicase_C; I.
SVART; SW00487; DENGC; I.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Created)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              01-MAY-2000 (TrEMBLrel. 13, 01-MAY-2000 (TrEMBLrel. 13, 01-MAK-2003 (TrEMBLrel. 23,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Virology 262:250-263(1999)
                                                                                                                                                                                                                                                                                                           Ouery Match
Best Local Similarity 82.4
Matches 168; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Genome polyprotein. Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             [1]
SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1129
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      174
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Bukh J.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       60Md60
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                RESULT 9
      ò
                                                                                                                                                                                                                                                                                                                                                                                 q
                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                              q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
```

```
1129 CGSSDLYLVTRHADVIPVRRRCDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGLFRAAV 1188
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ATQTFLATSINGVLWTVYHGAGTRTIASPKGPVTQMYTNVDKDLVGWQAPQGSRSLTPCT 113
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              114 CGSSDLYLVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAV 173
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 3 KKGSVVIVGRIN-----LSGDTAYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVST
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SMART; SMO487; DEXDC; 1.
PROSITE: PS00392; DDC_GAD_HDC_YDC; 1.
PROSITE: PS0507; RNRP_POSITIVE; 1.
AFPOSITE: PS50521; RDRP_VIRAL; 1.
ATP-binding; Coat protein; Envelope protein; Glycoprotein; Helicase; Hydrolase; Monstructural protein; Polyprotein;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Viruses; sarwa positive-strand viruses, no DNA stage; Flaviviridae; Hepacivirus.
NCBL_TaxID=11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               83.7%; Score 851.5; DB 12; Length 3015;
ilarity 82.4%; Pred. No. 7.7e-74;
Conservative 10; Mismatches 17; Indels 9;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Hydrolase; Nonstructural protein; roiffrocin. RNA-directed RNA polymerase; Transferase; Transmembrane.centioner 3015 AA; 328084 MW; E309F6318067D6CD CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     01-DEC-2001 (TrEMBLrel. 19, Created)
01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
01-DAR-2003 (TrEMBLrel. 23, Last annotation update)
NS3 proclease (Fragment).
Hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       181 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CTRGVAKAVDFIPVENLGTTMRSP 1212
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             174 STRGVAKAVDFIPVESLETTMRSP 197
                                                                                                                                                                                                                            InterPro; IPR002518; HCV_NS2.
InterPro; IPR004109; HCV_NS3.
InterPro; IPR004109; HCV_NS4a.
InterPro; IPR001049; HCV_NS4b.
InterPro; IPR002168; HCV_NS5a.
InterPro; IPR002168; HCV_RGRP.
InterPro; IPR001509; Hellcase_C.
InterPro; IPR001050; Hellcase_C.
InterPro; IPR001050; RNA_Pol_DS_PS.
InterPro; IPR007094; RNA_Pol_DS_PS.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        pfam; PF01542; HCV_capsid; 1. Pfam; PF01542; HCV_capsid; 1. Pfam; PF01542; HCV_capsid; 1. Pfam; PF01542; HCV_capsid; 1. Pfam; PF01540; HCV_NS2; 1. Pfam; PF01500; HCV_NS2; 1. Pfam; PF01000; HCV_NS4b; 1. Pfam; PF01000; HCV_NS4b; 1. Pfam; PF01500; HCV_NS4b; 1. Pfam; PF00271; helicase_C; 1. Pfam; PF00998; viral_RARP; 1. Pr0Dom; PP016652; HCV_NS1; 1. Pr0Dom; PP016665; HCV_NS1; 1. Pr0Dom; PP018665; HCV_NS1; 1.
EMBL; AF177039; AAF01181.1;
EMBL; AF177037; AAF01179.1;
                                                                                    InterPro; IPR001410; DEAD.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    PRELIMINARY;
                                                                                                                                                               Interpro: IPR002519; H
Interpro: IPR002511; H
Interpro: IPR002518; H
Interpro: IPR004109; HC
Interpro: IPR00145; H
Interpro: IPR001490; HC
Interpro: IPR001490; HC
Interpro: IPR00166; HC
Interpro: IPR00166; HC
Interpro: IPR001650; HC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Local Similarity
hes 168; Conserv
                                                               P27958; 1HEI
```

```
79 IASPKGPVTQMYTNVDKDLVGWQAPQGSRSLTPCTCGSSDLYLVTRHADVIPVRRKGDSR 138
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          64 IASPRGPVIQMYINVDQDLVGWPAPQGARSLIPCTCGSSDLYLYIRHADVIPVRRRGDSR 123
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           19 TAYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVSTATQTFLATSINGVLWTVYHGAGTRT 78
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            4 TAYAQOTRGLIGGIXTSLIGRDKNQVEGEVQIVSTAAQIFLATCINGVCWTVYBGAGTRI 63
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      139 GSLLSPRPISTLKGSSGGPLLCPAGHAVGIFRAAVSTRGVAKAVDFIPVESLETTWRS 196
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       124 GSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVCTRGVAKAVDFIPVENLETTMRS 181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Holland-Staley C.A., Kovari L.C., Golenberg E., Mayers D.L.;
"Genetic Diversity and response to IFN of the NS3 Protease Gene from Clinical Strains of the Hepatitis C Virus.";
Submitted (Apr.-2001) to the EMBL/GenBank/DDBJ databases.
EMBL; 78569241; ARS454566.1;
PinterPro; IPR004109; HCV_NS3;
Pfam; PF02907; HCV_NS3; 1.
                                                                                                                                                                                                                                                                    STRAIN-Pt.30;
Holland-Staley C.A., Kovari L.C., Golenberg E., Mayers D.L.;
Holland-Staley C.A., Rovari L.C., Golenberg E., Mayers D.L.;
Genetic Diversity and response to IFN of the NS3 Protease Gene from
Clinical Strains of the Hepatitis C Virus.";
Submitted (AFR-2011) to the EMBL/GenBank/DDBJ databases.
EMBL; AF36238; ARS456631;
Interpro; IPR004109; HCV_NS3:
Pfam; PF02907; HCV_NS3: 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
Hepacivirus.
                                                                                                                                                                                       Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Length 181;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             11; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        181 181
181 AA; 19123 MW; 1CAE817345ED809D CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                             1 1
181 181
181 AA; 19084 MW; 3B5E8161F2100A72 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                01-DEC-2001 (TrEMBLrel. 19, Created)
01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
                                                                                                                            Last sequence update)
Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                            83.3%; Score 847; DB 12; 92.7%; Pred. No. 4.9e-75; ative 2; Mismatches 11;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   181 AA.
                                                                                                              Created)
                                                                                                           01-DEC-2001 (TrEMBLrel. 19,
01-DEC-2001 (TrEMBLrel. 19,
01-MAR-2003 (TrEMBLrel. 23,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             165; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               NS3 protease (Fragment).
Hepatitis C virus.
                                                                                                                                                            protease (Fragment).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Similarity
                                                                                                                                                                                                                                         [1]
SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             NCBI_TaxID-11103;
                                                                                                                                                                         Hepatitis C virus
                                                                                                                                                                                                                        NCBI_TaxID=11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             STRAIN-Pt. 4V;
                                                                                                                                                                                                         Hepacivirus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          NON_TER
NON_TER
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                Protease.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Protease.
                                                                                                                                                                                                                                                                                                                                                                                                                                             SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Local
                                               RESULT 12
Q91RR5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   RESULT 13
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      윱
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ó
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    a
                                                                                                                                                                                                                                                                     ô
                                                                                                                                                                                                                                                                                                                                                              79 IASPKGPVTQMYTNVDKDLVGWQAPQGSRSLTPCTCGSSDLYLVTRHADVIPVRRRGDSR 138
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          64 IASPKGPVIQMYINVDKDLVGWPAPQGSRSLTPCTCGSSDLYLVTRHADVIPVRRGDSR 123
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           4 TAYAQQTRGLLGCIITSLIGRDKNQVEGEYQIVSTAAQTFLATCINGVCWTVYHGAGTRT 63
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          19 TAYAQQTRGEEGCQETSQTGRDKNOVEGEVQIVSTATQTFLATSINGVLWTVYHGAGTRT 78
                                                                                                                                                                                                                                                                                                     19 TAYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVSTATQTFLATSINGVLWTVYHGAGTRT 78
                                                                                                                                                                                                                                                                                                                      4 TAYAQOTRGLLGCIITSLTGRDKNQVEGEVQIVSTAAQTFLATCINGVCWTVYHGAGTRI 63
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          GSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVSTRGVAKAVDFIPVESLETIMRS 196
                                                                                                                                                                                                                                                                                                                                                                                                                              139 GSLLSPRPISTLKGSSGGPLLCPAGHAVGIFRAAVSTRGVAKAVDFIPVESLETTMRS 196
                                                                                                                                                                                                                                                                                                                                                                                                                                               124 GSLLSPRDISYLKGSSGGPLLCPAGHAVGIFRAAVCTRGVAKAVDFIPVENLETTMRS 181
                     "Canada Staley C.A., Kovari L.C., Golenberg E., Mayers D.L.;
"Cenetic Diversity and response to IFN of the NS3 Protease Gene from Clinical Strains of the Hepatitis C Virus.";
Submitted (APR-2001) to the EMBL/GenBank/DDBJ databases.
EMBL: AF369235; AAK54560.1; --
InterPro: IPR004109; HCV_NS3.
Pfam; PP02907; HCV_NS3.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Holland-Staley C.A., Kovari L.C., Golenberg E., Mayers D.L.;
"Genetic Diversity and response to IFN of the N33 Protease Gene from
"Genetic Diversity and response to IFN of the N33 Protease Gene from
Clinical Strains of the Hepatitis C Virus.";
Submitted (APR-2001) to the EMBL/GenBank/DDBJ databases.
EMBL, AF856218, AAR5443.1;
InterPro: IPR004109; HCV_NS3.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Gaps
                                                                                                                                                                                                                                                                         Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ;
0
                                                                                                                                                                                                                                                                     .;
                                                                                                                                                                                                                                  Score 849; DB 12; Length 181;
Pred. No. 3.1e-75;
1; Mismatches 11; Indels (
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Length 181;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          11; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           181 181
181 AA; 19130 MW; 85D91869299B7C35 CRC64;
                                                                                                                                                                         1 1
181 181
181 AA: 19130 MW; 85D91869299B7C35 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      01-DEC-2001 (TrEMBLrel. 19, Created)
01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
01-DEC-2003 (TrEMBLrel. 23, Last annotation update)
NS3 protease (Fragment).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Score 849; DB 12;
Pred. No. 3.1e-75;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           181 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1: Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            83.5%;
93.3%;
                                                                                                                                                                                                                                       83.5%;
93.3%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          166; Conservative
                                                                                                                                                                                                                                                                     Matches 166; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Similarity
                                                                                                                                                                                                                                                        Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SEQUENCE FROM N.A.
[1]
SEQUENCE FROM N.A.
STRAIN-Pt.1Y;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       NCBI_TaxID=11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    STRAIN-Pt.4;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Hepatitis C
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Hepacivirus
                                                                                                                                                                         NON_TER
NON_TER
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              NON_TER
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          139
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Protease.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Query Match
Best Local S
                                                                                                                                                                                                                                         Match
                                                                                                                                                                                                                                         Query Match
Best Local
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Q91RT5
Q91RT5;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             RESULT 11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        셤
 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                           셤
                                                                                                                                                                                                                                                                                                     à
                                                                                                                                                                                                                                                                                                                                                                  ò
```

Gaps

0;

```
Holland-Staley C.A., Kovari L.C., Golenberg E., Mayers D.L.;
"Genetic Diversity and response to IrN of the NS3 Protease Gene from
"Genetic Diversity and response to IrN of the NS3 Protease Gene from
Clinical Strains of the Hepatitis C Virus.";
Submitted (APR-2001) to the EMBL/GenBank/DDBJ databases.
EMBL, AF86540; ARK$465.1;
InterPro; IPR004109; HCV_NS3.
Pfam; PF02907; HCV_NS3.
                                                        SEQUENCE FROM N.A.
NCBI_TaxID-11103;
                                                                                                                                                                                                                                                                                                                              Protease.
ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       đ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                       ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                64 IASPKGPVIQMYINVDKDLVGWPAPQGARSLTPCTCGSSDLYLVTRHADVIPVRRKGDSR 123
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 79 IASPKGPVTQMYTNVDKDLVGWQAPQGSRSLTPCTCGSSDLYLVTRHADVIPVRRRGDSR 138
                                                                                                                                                                                                                                                                                                  64 IASPKGPVIQMYINVDQDLVGWPAPQGARSLTPCTCGSSDLYLVTRHADVIPVRRRGDSR 123
                                                                                                                                                                                                                                                                   79 IASPKGPVTOMYTNVDKDLVGWQAPQGSRSLTPCTCGSSOLYLVTRHADVIPVKRRGDSR 138
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    19 TAYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVSTATQTFLATSINGVLWTVYHGAGTRT 78
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   4 TAYAQQTRGLLGCIITSLTGRDKNQVEGEVQIVSTATQTFLATCINGVCWTVYHGAGTRT 63
                                                                                                                                                                                  4 TATAQQTRGLLGCIXTSLTGRDKNQVEGEVQIVSTAXQTFLATCINGVCWTVYHGAGTRT 63
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        GSLLSPRPISYLKGSSGGPI.LCPAGHAVGIFRAAVSTRGVARAVDFIPVESLETTMRS 196
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    139 GSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVSTRGVAKAVDFIPVESLETTMRS 196
                                                                                                                                                                                                                                                                                                                                                                                                                  19 TAYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVSTATQTFLATSINGVLWTVYHGAGTRT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      STRAIN-F.1C;
Holland-Staley C.A., Kovari L.C., Golenberg E., Mayers D.L.;
Holland-Staley C.A., Kovari L.C., Golenberg E., Mayers D.L.;
Genetic Diversity and response to Then of the N3 Protease Gene from Clinical Strains of the Hepatitis C Virus.";
Submitted (APR-2001) to the EMBL/GenBank/DDBJ databases.
EMBL; AF36914, AAK5439.1;
InterPro; IPR004109; HCV_NS3.
Pfam; PF02907; HCV_NS3.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Gaps
                                                                                       Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           NS3 protease (Fragment).
Hepatitis C virus.
Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Q91RR3 PRELIMINARY; PRT; 181 AA.
01-DEC-2001 (TrEMBLrel. 19, Created)
01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
01-DEC-2003 (TrEMBLrel. 23, Last annotation update)
NS3 protease (Fragment).
HS2 protease (Fragment).
Hepatitis C virus.
Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ;
                                                                                       .;
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                83,3%; Score 847; DB 12; Length 181; 92.7%; Pred. No. 4.9e-75;
                               Length 181;
                                                                                       11; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1 1
181 181
181 AA; 19114 MW; BEIDOBS42F014E86 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                01-DEC-2001 (TrEMBLrel. 19, Created)
01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
                     Score 847; DB 12;
Pred. No. 4.9e-75;
2; Mismatches 11;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         2; Mismatches
                        83.3%;
92.7%;
                                                                                    165; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Matches 165; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              PRELIMINARY;
                                                           Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           NCBI_TaxID-11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Hepacivirus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Hepacivirus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             NON_TER
NON_TER
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Query Match
Best Local S
                               Query Match
Best Local
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Q91RT9
Q91RT9;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   RESULT 14
091RT9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           RESULT 15
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 091RR3
1D 099
DT 019
DT 010
DT 010
DT 010
DE NS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 AND THE STATE OF T
                                                                                                                                                     δ
                                                                                                                                                                                                           q
                                                                                                                                                                                                                                                                      ö
                                                                                                                                                                                                                                                                                                                        ద
                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                               a
```

```
ö
                                                                                                                                                                                                                                                            64 IASPKGPVIQMYINVDRDLVGWPAPQGARSLTPCTGGSSDLYLVTRHADVIPVRRGDSR 123
                                                                                                                                                                                                                                        79 IASPRGPVTQMYTNVDKDLVGWQAPQGSRSLTPCTCGSSDLYLVTRHADVIPVRRRGDSR 138
                                                                                                                                                                         4 TAYAQOTRGLIGCIITSLIGRDKNQVEGEVQIVSTAAQTFLATCINGVCWTVYHGAGTRI 63
                                                                                                                                                                                                                                                                                                                         139 GSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVSTRGVAKAVDFIPVESLETTMRS 196
                                                                                                                                                                                                                                                                                                                                           19 TAYAQQIRGEEGCQETSQTGRDKNQVEGEVQIVSTATQIFLAISINGVLWTVYHGAGIRI
                                                                                                            Gaps
                                                                                                            ö
                                                                 Length 181;
                                                                                                          Indels
  181 181
181 AA; 19115 MW; 5D85F88AD7AC1A11 CRC64;
                                                            83.2%; Score 846; DB 12; I 92.7%; Pred. No. 6.1e-75; tive 2; Mismatches 11;
                                                                                                                                                                                                                                                                                                                                                                                                                             Search completed: August 30, 2003, 19:18:20 Job time : 38.5921 secs
                                                                                                          Conservative
                                                                                     Similarity
                                                                                                       165;
NON_TER
SEQUENCE
                                                                 Query Match
```

```
AR145264 Sequence
AR145263 Sequence
AR145263 Sequence
AR145263 Sequence
AR145265 Sequence
AR145267 Sequence
AR145269 Sequence
AR145269 Sequence
AR145263 Sequence
AR145253 Sequence
AR145253 Sequence
AR145274 Sequence
AR145274 Sequence
AR145275 Sequence
AR145255 Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    AF009606 Hepatitis
AR110831 Sequence
BD069985 Functiona
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        109331 Sequence 15
M32084 Hepatitis C
AR118703 Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           AR118686 Sequence
106434 Sequence 48
109328 Sequence 8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  AR118692 Sequence
106440 Sequence 54
109329 Sequence 10
                                                                                                                                                                                                                                                                                                  AR179057 Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             AR118696 Sequence
                                                                                                                                                                           Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.
                                                                                                                                                                                                                                                                           Description
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ALIGNMENTS
                                                                                                                                                                                                                                 SUMMARIES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     BD069982
AF009606
AR110831
BD069985
AR118686
                                                                                                                                                                                                                                                                                                                           AR145268
AR145262
AR145263
AR145254
AR145266
AR145267
AR145261
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ARI45270
ARI45274
ARI45272
ARI45256
ARI45257
ARI45260
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                AR145273
AR145271
AR145255
AR127810
BD081911
AR127809
BD081910
                                                                                                                                                                                                                                                                                                                                                                                                                                                 AR145258
AR145252
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  AR118692
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           AR118696
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           AR110828
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          AR14
                                                                                                                                                      em_htgo_other:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                             AR1
 em_htg_hum:*
em_htg_inv:*
em_htg_other:*
em_htg_mus:*
                                                                                                                         em_htgo_hum:*
                                                         em_htg_pln:*
em_htg_rod:*
                                                                                                 em_htg_vrt:*
                                                                                    em_htg_mam:*
                                                                                                                                                                                                                                                                         В
                                                                                                                                                                                                                                                           Query
Match Length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                12980
12980
5360
5360
                                                                                                                                                                                                                                                                                                RESULT 1
                                                                                                                                                                                                                                                                                                                                                                                                                       Result
                                                                                                                                                                                                                                                                         Š
                                                                                                                                                                                                                                                                                                                                                                                                                COMMAIND line parameters:
-WODEL-frame+ p2n.model -DEV-x1p
-WODEL-frame+ p2n.model -DEV-x1p
-WODEL-frame+ p2n.model -DEV-x1p
-D6-/cgn2_1/USPTO_spool/US0996594/runat_29082003_151919_28310/app_query.fasta_1.2872
-D8-GenEmb1 -QFNT-fastap -SUFFX-rye -MIRMATCH-0.1 -LOOPCL-0 -LOOPEXT-0
-DB-GenEmb1 -QFNT-fastap -SUFFX-rye -MIRMATCH-0.1 -LOOPCL-0 -LOOPEXT-0
-DOCALIGN-200 -THR_SCORE-pct -THR_MAX.** - DOCALIGN-15 -MODE-LOCAL
-DOCALIGN-200 -THR_SCORE-pct -THR_MAX.** - NALEN-2000000000
-USFR-050996594_GCGN1_1_14680_grunat_29082003_151919_28310_NCPU-6 -ICPU-3
-NO_MAAP -LARGEQUERY -NEG_SCORE-0 -WAIT -DSPBLOCK-100 -LONGLOG
-DEV_TIMEOUT-120 -WARN_TIMEOUT-30 -THREADS-1 -XGAPOP-10 -XGAPEXT-0.5 -FGAPOP-6
-FGAPEXT-7 -YGAPOP-10 -YGAPEXT-0.5 -DELOP-6 -DELEXT-7
                                                                          August 30, 2003, 19:18:33 ; Search time 2560.57 Seconds (without alignments) 3147.423 Million cell updates/sec
                                                                                                                                                             1 MKKKGSVVIVGRINLSGDTA.....VAKAVDFIPVESLETTMRSP 197
GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.
                                                       using frame_plus_p2n model
                                                                                                                                                                                                                                                                                               Total number of hits satisfying chosen parameters:
                                                                                                                                                                                                                                                                    2888711 seqs, 20454813386 residues
                                                                                                                                                                                                                                                                                                                                                               Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries
                                                                                                                                                                                                  Xgapop 10.0 , Xgapext
Ygapop 10.0 , Ygapext
Fgapop 6.0 , Fgapext
Delop 6.0 , Delext
                                                                                                                                                                                                                                                                                                                         Minimum DB seq length: 0
Maximum DB seq length: 2000000000
                                                                                                                             US-09-965-594-18
1017
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           gb_ba:*
gb_htg:*
gb_in:*
gb_ov:*
gb_ov:*
gb_pt:*
gb_pl:*
gb_pl:*
gb_pr:*
gb_sts:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                em_hum:*
em_in:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            em_ph:*
em_pl:*
em_ro:*
em_sts:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  em_pat:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          gb_un:*
gb_vi:*
em_ba:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     em_fun:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           em_mu:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       em_or:*
                                                       nucleic search,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   GenEmbl:*
                                                                                                                                                                                      BLOSUM62
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Title:
Perfect score:
                                                                                                                                                                                        Scoring table:
                                                       OM protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Database :
                                                                                                                                                              Sequence:
                                                                                                                                                                                                                                                                     Searched:
                                                                                 Run on:
```

us-09-965-594-18.rge

```
Percent Similarity:
Best Local Similarity:
Query Match:
DB:
AR145264
AR145264.1
                                                                                                                             ø
                           Unknown
                                     Unknown
                                                                                                                                                       Alignment Scores:
                                                                                                                                                                                                                                               S
                                                                                                                                                                                                                                                                64
                                                                                                                                                                                                                                                                                                                                                                                           82
        VERSION
KEYWORDS
SOURCE
ORGANISM
                                                                                                            source
                                                                                                                            BASE COUNT
ORIGIN
                                                                                                                                                                No.:
                                                     REFERENCE
AUTHORS
                                                                                         JOURNAL
ACCESSION
                                                                                                  FEATURES
                                                                        TITLE
                                                                                                                                                                         Score:
                                                                                                                                                                                                                                                                                                                                                                                                                                           q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                              q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                             qq
                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                 đ
                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                     οy
                                                                                                                                                                                                                                                                                                                                                                      qq
                                                                                                                                                                                                                                                                                                                                                                                          οy
                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                           οy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
         20-APR-2002
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             AlaGlyHisAlaValGlyIlePheArgAlaAlaValSerThrArgGlyValAlaLySAla 181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          952
                                                                                                                                                                                                                                                                                                                                                       592
                                                                                                                                                                                                                                                                                                                                                                                          652
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    LeuSerProArgProlleSerTyrLeuLysGlySerSerGlyGlyProLeuLeuCysPro 161
                                                                                                                                                                                                                                                                                413 GGTAGTGTTGTTGTTGGTAGAATTGTTTATCTGGTAGTGGTAGTATCACGGCGTAC 472
                                                                                                                                                                                                                                                                                                                                                                                                            ProLysGlyProValThrGlnMetTyrThrAsnValAspLysAspLeuValGlyTrpGln 101
                                                                                                                                                                                                                                                                                                  41
                                                                                                                                                                                                                                                                                                                                      61
                                                                                                                                                                                                                                                                                                                                                                        81
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 593 TGCATCAATGGGCTATGCTGGACTGTCTACCACGGGGCCGGAACGAGGACCATGGCATCA
                                                                                                                                                                                                                                                                                                                                      AsnGlnValGluGlyGluValGlnIleValSerThrAlaThrGlnThrPheLeuAlaThr
                                                                                                                                                                                                                                                                                                                                              SerIleAsnGlyValLeuTrpThrValTyrHisGlyAlaGlyThrArgThr1leAlaSer
                                                                                                                                                                                                                                                               GlyServalvalilevalGlyArglleAsnLeuSerGlyAsp------ThralaTyr
                                                                                                                                                                                                                                                                                                AlaGlnGlnThrArgGlyGluGluGlyCysGlnGluThrSerGlnThrGlyArgAspLys
                                                                      Unclassified.

1 (bases 1 to 12734)
Hong, Z., Lal, V.C.H. and Lau, J.Y.N.
Hepatitis C virus protease-dependent chimeric pestivirus
Patent: US 6326137-A 1 04-DEC-2001;
         PAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ValAspPheIleProValGluSerLeuGluThrThrMetArgSer 196
         linear
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Length:
Matches:
Conservative:
Mismatches:
Indels:
Gaps:
         DNA
                                                                                                                                               2803
     Sequence 1 from patent US 6326137.
AR179057
                                                                                                                                                                                                                                              US-09-965-594-18 (1-197) x AR179057 (1-12734)
                                                                                                                                  /organism="unknown"
2604 c 3295 g
                                                                                                                    Location/Oualifiers
                                   AR179057.1 GI:20220612
                                                                                                                                                                                1.11e-65
892.50
92.82%
90.77%
87.76%
                                                                                                                                                                                                Percent Similarity:
Best Local Similarity:
Query Match:
                                                                                                                                             4032 a
                                                      Unknown
                                                                Unknown
                                                                                                                                                                         Alignment Scores:
Pred. No.:
                                                                                                                                                                                                                                                                                                  22
                                                                                                                                                                                                                                                                                                                   473
                                                                                                                                                                                                                                                                                                                                                      533
                                                                                                                                                                                                                                                                                                                                                                                                                            623
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 773
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        162
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        893
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          182
                                                                                                                                                                                                                                                                S
                                                                                                                                                                                                                                                                                                                                     42
                                                                                                                                                                                                                                                                                                                                                                        62
                                                                                                                                                                                                                                                                                                                                                                                                                                               102
                                                                                                                                                                                                                                                                                                                                                                                                                                                               713
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                122
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  142
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    833
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            953
                                                                                                                                                                                                                                                                                                                                                                                                           82
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             RESULT 2
AR145264
LOCUS
DEFINITION
                                                                                                                              source
                 DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
                                                                                                                                             BASE COUNT
ORIGIN
                                                               ORGANISM
                                                                                                  TITLE
JOURNAL
                                                                                          AUTHORS
                                                                                REFERENCE
AR179057
LOCUS
                                                                                                                    FEATURES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            a
                                                                                                                                                                                                                             DB
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        d
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                               à
                                                                                                                                                                                                                                                                               a
                                                                                                                                                                                                                                                                                                  ô
                                                                                                                                                                                                                                                                                                                  QQ
                                                                                                                                                                                                                                                                                                                                     ç
                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                        à
                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                           a
                                                                                                                                                                                                                                                                                                                                                                                                                                               ç
                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ö
```

PAT 08-AUG-2001

linear

DIA

AR145264 1998 bp C Sequence 105 from patent US 6211338.

```
08-AUG-2001
                                                                                                                                                                                                                                                                                                                                                       243
                                                                                                                                                                                                                                                                                                                                                                        363
                                                                                                                                                                                                                                                                                                                                                                                                                                                     141
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              483
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    543
                                                                                                                                                                                                                                                                 183
                                                                                                                                                                                                                                                                                                                                                                                                                 101
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  161
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    47
                                                                                                                                                                                                                                                                                                                                    61
                                                                                                                                                                                                                                                         21
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                42 AsnGlnValGluGlyGluValGlnIleValSerThrAlaThrGlnThrPheLeuAlaThr
                                                                                                                                                                                                                                                                                                                                             ProLysGlyProValThrGlnMetTyrThrAsnValAspLysAspLeuValGlyTrpGln
                                                                                                                                                                                                                                                                                              22 AlaGlnGlnThrArgGlyGluGluGlyCysGlnGluThrSerGlnThrGlyArgAspLys
                                                                    NS3
                                    Unclassified.

1 (bases 1 to 1998)
Malcolm,B.A., Taremi,S.Shane., Weber,P.C. and Yao,N.
Single-chain recombinant complexes of hepatitis C virus protease and NS4A cofactor peptide
Patent: US 6211338-A 105 03-APR-2001;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         PAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           linear
                                                                                                                                                                                                                                                         GlySerValValIleValGlyArgIleAsnLeuSerGlyAsp---
                                                                                                                                                                  1998
167
16
10
3
                                                                                                                                                                  Length:
Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         DNA
                                                                                                                             u
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      AR145268 1998 bp D)
Sequence 109 from patent US 6211338.
AR145268.1 GI:15107135
                                                                                                                                                                                                                    Gaps:
                                                                                                                                                                                                                                     US-09-965-594-18 (1-197) x AR145264 (1-1998)
                                                                                               Location/Qualifiers
1. .1998
                                                                                                                   /organism~"unknown"
595 c 569 q
                                                                                                                             569 9
GI:15107131
                                                                                                                                                                  1.1e-65
881.50
93.37%
85.20%
86.68%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    RESULT 3
AR145268
LOCUS
DEFINITION
ACCESSION
VERSION
```

```
Unclassified
Unknown
                                                                                                                                                   Query Match:
                                                                 source
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         DEFINITION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
                                                                                 BASE COUNT
                                                                                                                   ..
Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ORGANISM
                        AUTHORS
TITLE
                                                JOURNAL
                REFERENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          AR145263
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 RESULT 5
                                                         FEATURES
                                                                                         ORIGIN
                                                                                                                                                                                          ò
                                                                                                                                                                                                           a
                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                            a
                                                                                                                                                                                                                                                                                                            a
                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                             a
                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                             셤
                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                              ద
                                                                                                                                                                                                                                                                                                                                                                                                                                                               õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                               셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               셤
                                                                                                                                                                                                                           à
                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                            ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  08-AUG-2001
                                                                                                                                                                                                                                          243
                                                                                                                                                                                                                                                                                                                     122 ValThrArgHisAlaAspValIleProValArgArgArgGlyAspSerArgGlySerLeu 141
                                                                                                                                                                                                                                                                                                                                                                                                                      LeuSerProArgProlleSerTyrLeuLysGlySerSerGlyGlyProLeuLeuCysPro 161
                                                                                                                                                                                                                                                                                                                                                                                                                                                                              162 AlaGlyHisAlaValGlyIlePheArgAlaAlaValSerThrArgGlyValAlaLysAla 181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               603
                                                                                                                                                                                                                                                                            61
                                                                                                                                                                                                                                                                                                            81
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CTCTCCCCCAGGCCTGTCTCTACTTGAAGGCTCTGCTGGTGGTCCACTGCTCTGCCCT
                                                                                                                                                                                                           GlySerValValIleValGlyArgIleAsnLeuSerGlyAsp-----ThrAlaTyr
                                                                                                                                                                                                                                                                                                            SerIleAsnGlyValLeuTrpThrValTyrHisGlyAlaGlyThrArgThrIleAlaSer
                              1 (bases 1 to 1998)
Malcolm,B.A., Taremi,S.Shane., Weber,P.C. and Yao,N.
Single-chain recombinant complexes of hepatitis C virus NS3
profease and NS4A cofactor peptide
Patent: US 6211338-A 109 03-APR-2001;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  PAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ValAspPheIleProValGluSerLeuGluThrThrMetArgSerPro 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        linear
                                                                                                                                 1998
166
17
10
3
                                                                                                                                Length:
Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Sequence 103 from patent US 6211338.
AR145262
                                                                                                                                                                           Gaps:
                                                                                                                                                                                          US-09-965-594-18 (1-197) x AR145268 (1-1998)
                                                                      Location/Qualifiers
1. .1998
                                                                                         /organism="unknown"
595 c 569 g
                                                                                                  Þ
                                                                                                569
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          AR145262.1 GI:15107129
                                                                                                                                 1.99e-65
878.50
93.37%
84.69%
86.38%
                        Unclassified
                                                                                                                                                 Percent Similarity:
Best Local Similarity:
                                                                                                 40
                Unknown
                                                                                                                         Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               182
                                                                                                                                                                                                                                                                                                                                                                                                                                             142
                                                                                                                                                                                                                                                                                                            62
                                                                                                                                                                                                                                                                                                                                            82
                                                                                                                                                                                                                                                                                                                                                            304
                                                                                                                                                                                                                                                                                                                                                                                            364
                                                                                                                                                                                                                                                                                                                                                                                                                                                              184
                                                                                                                                                        Best Local Si
Query Match:
DB:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
                                                                                 source
       SOURCE
                                                                                                BASE COUNT
                                                                                                                                  Pred. No.:
                                        AUTHORS
TITLE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                RESULT 4
AR145262
LOCUS
                                                                JOURNAL
                                REFERENCE
                                                                        FEATURES
                                                                                                         ORIGIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                          q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ô
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               a
                                                                                                                                                                                                                                                           용
                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                            οy
                                                                                                                                                                                                                                                                                                                                                                                                                            8
                                                                                                                                                                                                                                                                                                                                                                                                                                         6
                                                                                                                                                                                                           ö
                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                           ŏ
                                                                                                                                                                                                                                                                            Š
                                                                                                                                                                                                                                                                                                            ö
                                                                                                                                                                                                                                                                                                                        ద
                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                             õ
                                                                                                                                                                                                                                                                                                                                                                                        g
```

```
PAT 08-AUG-2001
                                                                                                                                                                                                                                                                                                               ValThrArgHisAlaAspValIleProValArgArgArgGlyAspSerArgGlySerLeu 141
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            102 AlaProGlnGlySerArgSerLeuThrProCysThrCysGlySerSerAspLeuTyrLeu 121
                                                                                                                                                                                                                                                                                                                                                         62 SerileAsnGlyValLeuTrpThrValTyrHisGlyAlaGlyThrArgThrileAlaSer 81
                                                                                                                                                                                                                                                                          41
                                                                                                                                                                                                                                         GlySerValValIleValGlyArglleAsnLeuSerGlyAsp------ThrAlaTyr
Unclassified.

1 (bases 1 to 1998)
Malcolm.B.A., Taremi,S.Shane., Weber,P.C. and Yao,N.
Single-chain recombinant complexes of hepatitis C virus ly protease and NS4A cofactor peptide
Patent: US 621138-A 103 03-APR-2001;
Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        linear
                                                                                                                                    1998
167
15
11
3
                                                                                                                                    Length:
Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Sequence 104 from patent US 6211338, AR145263
                                                                                             424
                                                                                                                                                                                                          (1-1998)
                                                                                  /organism="unknown"
596 c 568 g
                                                                                                                                                                                                           US-09-965-594-18 (1-197) x AR145262
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       GI:15107130
                                                                                                                                  2.42e-65
877.50
92.86%
85.20%
86.28%
                                                                         1. .1998
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      AR145263.1
                                                                                                                                                                    Best Local Similarity:
                                                                                             410 a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Unknown
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Unknown
                                                                                                                                                          Percent Similarity:
                                                                                                                            Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                122
                                                                                                                                                                                                                               2
                                                                                                                                                                                                                                                                                                                                                                                                 82
                                                                                                                                                                                                                                                                                                                                                                                                                   304
                                                                                                                                                                                                                                                                                                                                                                                                                                                           364
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   424
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              604
```

us-09-965-594-18.rge

```
Unclassified
                                                                                                       Percent Similarity:
Best Local Similarity:
                                                     120 a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Unknown
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Unknown
                                                                              Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                              142
                                                                                                                        Ouery Match:
DB:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
                                   source
                                                                                        ٠.
و
                                                     BASE COUNT
                   JOURNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         REFERENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  AUTHORS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          AR145266
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        TITLE
                            FEATURES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   RESULT
                                                              ORIGIN
                                                                                        Pred.
                                                                                                Score
                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                   qq
                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                        a
                                                                                                                                                                                                                                                                                                                                                                           5
                                                                                                                                                                                                                                                                                                                                                                                            셤
                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  음
                                                                                                                                                                   ð
                                                                                                                                                                                 g
                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                     à
                                                                                                                                                                                                                                                                     ŏ
                                                                                                                                                                                                                                                                                                                                           à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    PAT 08-AUG-2001
                                                                                                                                                                                                                                                                                                                                                                                          AlaProGlnGlySerArgSerLeuThrProCysThrCysGlySerSerAspLeuTyrLeu 121
                                                                                                                                                                                                   64 GGTTCTGTTGTTGTTGTTGTAGAATTATTTTATCTGGTAGTGGTAGTATCACGGCCTAC 123
                                                                                                                                                                                                                                                                                                                        ProLysGlyProValThrGlnMetTyrThrAsnValAspLysAspLeuValGlyTrpGln 101
                                                                                                                                                                                                                                                                                      SerIleAsnGlyValLeuTrpThrValTyrHisGlyAlaGlyThrArgThrIleAlaSer 81
                                                                                                                                                                                   GlySerValValIleValGlyArgIleAsnLeuSerGlyAsp-----ThrAlaTyr
1 (bases 1 to 1998)
Malcolm, B.A., Taremi, S. Shane., Weber, P.C. and Yao, N.
Single-chain recombinant complexes of hepatitis C virus NS3
protease and NS4A cofactor peptide
Patent: US 6211338-A 104 03-APR-2001;
Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1 (bases 1 to 651)
Malcolm, B.A., Taremi, S.Shane., Weber, P.C. and Yao, N.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     linear
                                                                                                      Length:
Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     DNA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              AR145254 651 bp
Sequence 95 from patent US 6211338.
AR145254
                                                                    424
                                                                                                                                                 Gaps:
                                                                                                                                                                 US-09-965-594-18 (1-197) x AR145263 (1-1998)
                                                           /organism="unknown"
596 c 568 q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            AR145254.1 GI:15107121
                                                                                                       2.42e-65
877.50
92.86%
85.20%
86.28%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Unclassified
                                                                                                                      Percent Similarity:
Best Local Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Unknown.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Unknown
                                                                                                Alignment Scores:
                                                                                                                                                                                                                                                                                      62
                                                                                                                                                                                                                                                                                                                                                         102
                                                                                                                                                                                                                                                                                                                                                                                                            424
                                                                                                                                                                                                                                                                                                                                                                                                                             142
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               604
                                                                                                                                                                                                                                                                                                                       82
                                                                                                                                                                                                                                                                                                                                                                                                                                             84
                                                                                                                               Best Local S
Query Match:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          AR145254
LOCUS
DEFINITION
ACCESSION
                                                      source
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ORGANISM
                                                                                                       Pred. No.:
                                                                    BASE COUNT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            VERSION
KEYWORDS
SOURCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       REFERENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                AUTHORS
                                  JOURNAL
         AUTHORS
                                            FEATURES
                  TITLE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   RESULT
                                                                              ORIGIN
                                                                                                                 Score:
                                                                                                                                                                                                                                                                                                  q
                                                                                                                                                                                                                                                                                                                                                                         q
                                                                                                                                                                                                                                                                                                                                                                                                         a
                                                                                                                                                                                                                                                                                                                                                                                                                                         a
                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                             đ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             셤
                                                                                                                                                                                   ó
                                                                                                                                                                                                   a
                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                    qq
                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                       ç
                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                          à
                                                                                                                                                                                                                                                                                                                                                                                                                             ö
```

```
PAT 08-AUG-2001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    363
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                141
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              162 AlaGlyHisAlaValGlyIlePheArgAlaAlaValSerThrArgGlyValAlaLysAla 181
                                                                                                                                                                                                                                                                                                                                                                                                                          243
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               101
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         LeuSerProArgProIleSerTyrLeuLysGlySerSerGlyGlyProLeuLeuCysPro 161
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            603
                                                                                                                                                                                                                                                                                               124 TCCCAACAGAGGGGGGGCCTACTTGGTTGCAAGAAGACTAGCCTTACAGGCCGGGACAAG 183
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 303
                                                                                                                                                                                                                                                                                                                                       41
                                                                                                                                                                                                                                                                                                                                                                                              61
                                                                                                                                                                                                                                                                                                                                                                                                                                                      83
                                                                                                                                                                                                                                                                                7
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  62 SerIleAsnGlyValLeuTrpThrValTyrHisGlyAlaGlyThrArgThrIleAlaSer
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           304 CCAAAGGGGCCAATCACCCAGATGTACACTAATGTGGACCAGGACCTCGTCGGCTGGCAG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             122 ValThrArgHisAlaAspValIleProValArgArgArgGlyAspSerArgGlySerLeu
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         GlySerValValIleValGlyArgIleAsnLeuSerGlyAsp-----ThrAlaTyr
                                                                                                                                                                                                                                                                                                                                                                                              42 AsnGlnValGluGlyGluValGlnIleValSerThrAlaThrGlnThrPheLeuAlaThr
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1 (Dasses 1 to 1998)
Malcolm, B.A., Taremi, S.Shane., Weber, P.C. and Yao, N.
Single-chain recombinant complexes of hepatitis C virus NS3
protease and NS4A cofactor peptide
   hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      182 ValAspPheIleProValGluSerLeuGluThrThrMetArgSer 196
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                linear
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      651
166
10
3
                                                                                                                                                Length:
Matches:
Conservative:
Mismatches:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              DNA
Single-chain recombinant complexes of protease and NSAA cofactor peptide Patent: US 6211338-A 95 03-APR-2001; Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         6211338.
                                                                                                                                                                                                       Indels:
Gaps:
                                                                                       144
                                                                                                                                                                                                                                                 (1-651)
                                                                     /organism-"unknown"
187 c 200 a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           AR145266 1998
Sequence 107 from patent US
AR145266
AR145266.1 GI:15107133
                                                                                                                                                                                                                                                  US-09-965-594-18 (1-197) x AR145254
                                                                                                                                             1.18e-65
874.50
93.33%
85.13%
85.99%
                                                        1. .651
```

1. .1998

	Patent: US 6211338-a 107 03-aPR-2001; Location/Qualifiers		source	
source	11998 /organism-"unknown" 10 = 506 = 569 = 424		BASE COUNT ORIGIN	
DASE COUNT	<b>5</b>		Alignment S	Scores
Alignment Scores Pred. No.:	: 4.35e-65 Length: 1		Score: Percent Similari	llari
Score: Percent Similarity: Best Local Similarity	## 150 Matches: 150 rity: 92.86% Conservative: 16		Best Local Query Match	Simil :
Query March: DB:	bo.yy* inders:		08-09-965-5	-594-18
US-09-965-594-18	18 (1-197) x AR145266 (1-1998)		Qy	5 G1
Qy 5	GlySerValValIleValGlyArgIleAsnLeuSerGlyAspThrAlaTyr	ThrAlaTyr 21	qu	64 99
Db 64 (		AGTATCACGCCTAC 123	ογ	22 A1
Oy 22	AlaGlnGlnThrArgGlyGluGluGlyCysGlnGluThrSerGlnThrGlyArgAspLys	PhrGlyArgAspLys 41	Db 1	124 TC
124	TCCCAACAGAGGGGGCTACTTGGTTGCAAGATCACTAGCCTT	ACAGGCCGGACAAG 183	Qy	42 As
42	AsnGlnValGluGlyGluValGlnIleValSerThrAlaThrGlnThrPheLeuAlaThr 			
184	aaccaggtcgagggagggttcaggtggtttccaccgcaacacaa	rcctrcctggcgAcc 243	Οÿ	62 Se
0y 62	SerIleAsnGlyValLeuTrpThrValTyrHisGlyAlaGlyThrArgThr1leAlaSer	ArgThrileAlaSer 81	op 9	244 TG
Db 244	tgcgtcaacgccgtgttggaccgtttaccatggtgctggctca	AAGACCTTAGCCGGC 303	οy	82 Pz
82	ProLysGlyProValThrGlnMetTyrThrAsnValAspLysAspLeuValGlyTrpGln	10	Dp 3	304 CC
Db 304	CCAAAGGGGCCAATCACCCAGATGTACACTAATGTGGACCAGGAC	CTCGTCGGCTGGCAG 363	Oy 1	102 A1
	AlaProGlnGlySerArgSerLeuThrProCysThrCysGlySerSerAspLeuTyrLeu	SerAspLeuTyrLeu 121	Dp 3	364 GC
Db 364	GCGCCCCCGGGGCGCGTTCCTTGACACCATGCACCTGTGGCAGC	FCAGACCTTTACTTG 423	0.7	122 Va
Oy 122	ValThrargHisalaAspVallleProValargArgArgGlyAspSerArgGlySerLeu	SerArgGlySerLeu 141	<b>4</b> qq	424 GT
Db 424	GTCACGAGACATGCTGACGTCATTCCGGTGGGCGGGGGGGG	AGTAGGGGAGCCTG 483	OY 1	142 Le
Qy 142	LeuSerProArgProlleSerTyrLeuLysGlySerSerGlyGlyProLeuLeuCysPro	ProLeuLeuCysPro 161	Db da	484 CT
Db 484	CTCTCCCCCAGGCCTGTCTCTTGAAGGGCTCTGCTGGTGGT	CCACTGCTCTGCCT 543	0y 1	162 Al
0y 162	AlaGlyHisAlaValGlyIlePheArgAlaAlaValSerThrArgGlyValAlaLySAla	SlyvalAlaLysAla 181	go qo	544 TC
Db 544	TCGGGGCACGCTGTGGGCATCTTCCGGGCTGCCGTATGCACCCGG	GGGTTGCGAAGGCG 603	oy 1	182 Va
Oy 182 Db 604	ValAspPhelleProValGluSerLeuGluThrThrMetArgSerPro	Pro 197       CG 651	Db 6	604 63
AR145267 LOCUS LOCUS DEFINITION SE ACCESSION AR VERSION REWORDS COURDS	AR142267 Sequence 108 from patent US 6211338. AR145267 GF:15107134	ear PAT 08-AUG-2001	AR145261 LOCUS DEFINITION ACCESSION VERSION KEYWORDS SOURCE ORGANISM	ARIA ARIA ARIA Unkr
	Onknown. Unclassified. 1 (bases 1 to 1998)		REFERENCE AUTHORS TITLE	Malc Malc Sing
AUTHORS MA TITLE SI JOURNAL PA FEATURES	Malcolm, B.A., Taremis, S.Dane, "Weber, P.C. and Tao, N. Single-chain recombinant complexes of hepatitis C virus protease and NS4A cofactor peptide Patent: US 621138-A 108 03-APR-2001; Location/Oualifiers	ac,n. C virus NS3	JOURNAL FEATURES SOURCE	Pate

```
PAT 08-AUG-2001
                                                                                                                                                                                                      243
                                                                                                                                                                          61
                                                                                                                                                                          spGlnValGluGlyGluValGlnIleValSerThrAlaThrGlnThrPheLeuAlaThr
                                                                                                                                                                                  (bases 1 to 1998)

loolm, B.A., Taremi, S.Shane, Weber, P.C. and Yao, N.

nofte-Chain recombinant complexes of hepatitis C virus NS3

otease and NS4A cofactor peptide

tent: US 6211338-A 102 03-APR-2001;

Location/Qualifiers

l. . . 1998

/organism="unknown"
                                                                                                                                                                                                                                                                                                                                                                                                1998
166
16
11
                                     Length:
Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                          1998 bp DNA equence 102 from patent US 6211338. 145261 GI:15107128
       424
                                                                             Gaps:
                                                                                           8 (1-197) x AR145267 (1-1998)
/organism="unknown"
596 c 568 g
                                    4.35e-65
874.50
92.86%
84.69%
85.99%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     lassified
                                                     ity:
|larity:
       410 a
```

```
VRL 12-JUL-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Pestivirus type 1
Pestivirus type 1
Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                   561
                                                                                                                                                                                                                                                                                                             121
                                                                                                                                                                                                                                                                                                                                 441
                                                                                                                                                                                                                                                                                                                                                                                             161
                                                                                                                                                                                                                                                                   101
                                                                                                                                                                                                                                                                                       381
                                                                                                                                                                                                                                                                                                                                                                                                                                     162 AlaGlyHisAlaValGlyIlePheArgAlaAlaValSerThrArgGlyValAlaLysAla 181
                                                                                                                                                              201
                                                                                                                                                                                                      261
                                                                                                                                                                                                                                               321
                                                                                                                     82 GGTTCTGTTGTTGTTGGTAGAATTATTTTATCTGGTAGTGGTAGTATCACGGCCTAC 141
                                                                                                                                          41
                                                                                                                                                                                  61
                                                                                                                                                                                                                           81
                                                                                                                                                      22 AlaGlnGlnThrArgGlyGluGluGlyCysGlnGluThrSerGlnThrGlyArgAspLys
                                                                                                                                                                                  42 AsnGlnValGluGlyGluValGlnIleValSerThrAlaThrGlnThrPheLeuAlaThr
                                                                                                                                                                                                                          AlaProGlnGlySerArgSerLeuThrProCysThrCysGlySerSerAspLeuTyrLeu
                                                                                                  GlyServalvalileValGlyArglleAsnLeuSerGlyAsp-----ThrAlaTyr
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Lal.V.C., Zhong.W., Skelton,A., Ingravallo,P., Vassilev,V., Donis,R.O., Hong.Z. and Lau.J.Y.
Donis,R.O., Hong.Z. and Lau.J.Y.
Donis,R.O., Hong.Z. and Lau.J.Y.
Donis,R.O., Hong.Z. and Lau.J.Y.
Droft-asse-dependent bovine viral diarrhea virus
J. Virol. 74 (14), 6339-6347 (2000)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                             182 ValAspPheIleProValGluSerLeuGluThrThrMetArgSerPro 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        linear
     2016
167
14
12
3
      Length:
Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          RNA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         AF268278 12734 bp
pestivius type 1, complete genome.
AF268278.1 GI:9049956
                                                          Gaps:
                                                                              US-09-965-594-18 (1-197) x AR145269 (1-2016)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        2 (bases 1 to 12734)
Lai, V.C.H. and Hong, Z.
Direct Submission
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       (bases 1 to 12734)
     5.35e-65
873.50
92.35%
85.20%
85.89%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Pestivirus.
                           Percent Similarity:
Best Local Similarity:
Query Match:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     20323484
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 10864644
Alignment Scores:
                                                                                                                                                                                                                            62
                                                                                                                                                                                                                                                                                        322
                                                                                                                                                                                                                                                                                                             102
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  622
                                                                                                                                                                                                                                                                                                                                 382
                                                                                                                                                                                                                                                                    82
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  DEFINITION
ACCESSION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        VERSION
KEYWORDS
SOURCE
ORGANISM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      MEDLINE
PUBMED
REFERENCE
AUTHORS
TITLE
        Pred. No.:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       REFERENCE
AUTHORS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      RESULT 11
AF268278
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           JOURNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       TITLE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           LOCUS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ద
                                                                                                                                                          g
                                                                                                                                                                                                g
                                                                                                                                                                                                                      õ
                                                                                                                                                                                                                                             q
                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                   a
                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                a
                                                                                                                                                                                                                                                                                                                                                Š
                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                     à
                                                                                                                                                                                                                                                                                                                                                                                                                                                          đ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Š
                                                                                                  õ
                                                                                                                   a
                                                                                                                                         ò
                                                                                                                                                                                 \delta
```

~

```
INDITIONS TO ANY INCORPOLATION OF THE MEAN AND THE ALTONOMY TO THE ALGORITHM AND THE ALGORITHM ALGORITHM ALGORITHM ALGORITHM AND THE ALGORITHM AND THE ALGORITHM ALGORITHM ALGORITHM ALGORITHM AND THE ALGORITHM A
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                /Cranslation-"MELMTNEGSGSVVIVGRIVLSGSGSITACAQQTRGLLGCKITSL
TGROKNQVEGEVQIVSTATOFFLATCINGVCMTVYHGAGTRTASPKGPVIOMYTNVD
ODLVGWPAPAGGSRSLIPOFTCASSDLIVTRHANUTPVRRGDSRGSLLSPRPISYLKG
SSGGPLLCPAGHAVGLERAAVCTRGVARAVDFIPVENLETTTRSGSGADTEDVVCCSM
SYSDTREGGATKKTQKPDRLERGKMKIVPKESEKDSKTRPPDATIVVEGVKYQVRKK
GKTKSKNTQDCLYTNKNKPQESRKKLEKALLAMAIIAIVLFOYTMGBNITGNNCDNG
TGGIORAAPORGVNSLLGGIPPERICTGVPSHLEADIELKTIHGMADASEKTNYTCCR
LQRHEBNKHGRCNWYNIEPWILVMNRTQANITEGOPPRECATTCRYDBASDLLWYTGA
RDSPTPLTGCKKGKNPSFAGILLANGCNFEIAASDVLFKEHERISWFQDTTLYLVDGL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   YILDLIYGLHKQINRGLKKWVLGWAPAPFSCDWTPSDERIRLPTDNYLRVETRCPCGY
EMKAFKNVGGKLTKVEESGPFLCRNRPGRGPVNYRVTKYYDDNLREIKPVAKLEGQVE
HYYKGVTAKIDYSKGKMLLATDKWEVEHGVITRLAKRYTGVGFNGAYLGDEPNHRALV
. (16-MAY-2000) Antiviral Therapy, Schering-Plough Research, , 2015 Galloping Hill Road, Kenilworth, NJ 07033-0539, USA Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       VLEKVEPDASDNQNSVKIGLDEGNYPGPGIQTHTLTEEIHNRDARPFIMILGSRNSIS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              NRAKTARNINLYTGNDPREIRDLMAAGRMLVVALRDVDPELSEMVDFKGTFLDREALE
ALSLGQPRPRQVTREAVRNLIEQKKDVEIPNWFASDDPVPLEVALKNDRYYLVGDVGE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ERDCATITKNTVOFLKMKKGCAFTYDLTISNLTRLIELVHRNNLEEKEIPTATVTTWL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       AYTEVNEDVGTIK PVLGERVIPDPVVDINLQPEVQVDTSEVGITIIGRETLMTTGVTP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          vkdoakalgatdotriikevgsrtyamklsswfloasnromsliplfeelllrcppat
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       TNSLEGAROGTAKLTTWLGKOLGILGKKLENKSKTWFGAYAASPYCDVDRKIGYIWYT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               KSNKGHMASAYQLAQGNWEPLGCGVHLGTIPARRVKIHPYEAYLKLKDFIEEEEKKPR
                                                                                                                                                                    /organism-"Pestivirus type 1"
/mol_type-"genomic RNA"
/db_xref-"taxon:11099"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  /product="polyprotein"
/protein_id="AAF82566.1"
/db_xref="G1:9049957"
                                                                                                                                                                                                                                                                                                                                                                                                                           /codon_start=1
                                                                                                                                                                                                                                                                                                                                                                 .12508
                                                                                                                                                                                                                                                                                                                                    386
                                                Institute,
Submitted
                                                                                                                                                          source
                                                                                                                                                                                                                                                                                                                               5'UTR
CDS
    JOURNAL
```

```
DTITDEMTEVPVITAGESTRINGORSGOPTSAGNSMLNUTHMYRECESSTGVPYK
SFNRYARIHVGEDGFLITERGLGIKFANKGMOILHEBAGFPOKITEGERMYRYRFED
IEFCSHTPPVFWSDNTSSHAGREDTAVILSKWATRLDSGGERGTTAVERAVAFSFLL
MYSWNELVRRICLLVISQOPETDPSKHATYYRKGDFTGAYKDVIGKRISELKRTGFER
IANLWISLSTLGIWTKHTSKRIIQDCVAIGKEBGNWLVNADRLISSTTGHLYIPDKGF
ILGGRIFFGQALDRTFETNRYMGYGTERYKLGPIVNLLLRRLKILLMTAVGVSS*

1 2608 c. 3293 g. 2802 t. 1 others
           YGEVTWEQLEAGINRKGAAGFLEKKNIGEVLDSERHLVEQLVRDLKAGRKIKYYETAI
PKNEKRDVSDDWQAGDLVVEKRPRVIQYPEAKTRLAITKVMYNWVKQOPVVIPGYEGK
TPLFNIFDKVRKEWDSFNEPVAVSFDTKAWDTQVTSKDLQLIGEIQKYYYKKEWHKFI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      08-AUG-2001
agirleklpivraqtdtktfheairdkidksenronpelhnklleifhtiaqptlkht
                                                                                                                                                                                                                                                                                                                                                                                                                                                   652
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              772
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           832
                                                                                                                                                                                                                                                                                                                                                      532
                                                                                                                                                                                                                                                                                                                                                                                                      592
                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ProLysGlyProValThrGlnMetTyrThrAsnValAspLysAspLeuValGlyTrpGln 101
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CCCAAGGGTCCTGTCATCCAGATGTATACCAATGTGGACCAAGACCTTGTGGGCTGGCCC 712
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   141
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   161
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        892
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        952
                                                                                                                                                                                                                                                                                                         GGTAGTGTTGTTGTTGGTAGAATTGTTTTATCTGGTAGTGGTAGTATCACGGCGTGC 472
                                                                                                                                                                                                                                                                                                                                                                                                                           81
                                                                                                                                                                                                                                                                                                                                                                              61
                                                                                                                                                                                                                                                                                                                                              GlySerValValileValGlyArgileAsnLeuSerGlyAsp-----ThrAlaTyr
                                                                                                                                                                                                                                                                                                                               22 AlaGlnGlnThrArgGlyGluGluGlyCysGlnGluThrSerGlnThrGlyArgAspLys
                                                                                                                                                                                                                                                                                                                                                                             AsnGlnValGluGlyGluValGlnIleValSerThrAlaThrGlnThrPheLeuAlaThr
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        AlaProGlnGlySerArgSerLeuThrProCysThrCysGlySerSerAspLeuTyrLeu
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               AlaGlyHisAlaValGlyIlePheArgAlaAlaValSerThrArgGlyValAlaLysAla
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      PAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ValAspPheIleProValGluSerLeuGluThrThrMetArgSer 196
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     997
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      linear
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Length:
Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       DNA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      651 bp
from patent US 6211338.
                                                                                                                                                                                                                                                            US-09-965-594-18 (1-197) x AF268278 (1-12734)
                                                                                                                                                                                                                                       Gaps:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       GI:15107125
                                                                                                                                                                        5.58e-64
872.50
91.79%
89.23%
85.79%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  AR145258
Sequence 99 f
AR145258
                                                                                                                                                                                              Percent Similarity:
Best Local Similarity:
Query Match:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       AR145258.1
                                                                                                                               ๙
                                                                                                                              4030
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Unknown.
                                                                                                                                                                 Alignment Scores:
                                                                                                                                                                                                                                                                                   ß
                                                                                                                                                                                                                                                                                                         113
                                                                                                                                                                                                                                                                                                                                                     473
                                                                                                                                                                                                                                                                                                                                                                              42
                                                                                                                                                                                                                                                                                                                                                                                                     533
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 653
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        102
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             713
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   122
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           773
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  142
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        833
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                162
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       893
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              182
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     953
                                                                                                                                                                                                                                                                                                                                                                                                                                                                         82
                                                                                                                 3'UTR
BASE COUNT
                                                                                                                                                                            Pred. No.:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              DEFINITION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ACCESSION
VERSION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             RESULT 12
AR145258
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 KEYWORDS
                                                                                                                                         ORIGIN
                                                                                                                                                                                         Score:
                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     셤
                                                                                                                                                                                                                                                                                   à
                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                            à
                                                                                                                                                                                                                                                                                                                                                                                              Q
                                                                                                                                                                                                                                                                                                                                                                                                                    õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             à
```

Unknown

ORGANISM

VKDTVIREHNKWILKKIRFQGNLNTKKMLNPGKLSEQLDREGRKRNIYNHQIGTIMSS

```
LOCUS
DEFINITION
                                                      BASE COUNT
ORIGIN
                                                                                     .. ON .
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ACCESSION
VERSION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ORGANISM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  AUTHORS
TITLE
AUTHORS
TITLE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                           RESULT 14
AR145253
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           REFERENCE
                       JOURNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           KEYWORDS
SOURCE
                               FEATURES
                                                                                             Score:
                                                                                                                                                                                                                                                                                                                                                                                                                             a
                                                                                                                                                                                                                                                                                                                               d
                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                              22
                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                  셤
                                                                                                                                                                                                                                                                                   ô
                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                          ô
                                                                                                                                                                      a
                                                                                                                                                                                        ολ
                                                                                                                                                                                                     셤
                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           08-AUG-2001
                                                                                                                                                                                                                                                                                                  243
                                                                                                                                                                                                                                                                                                                                                                 ValThrArgHisAlaAspValIleProValArgArgArgGlyAspSerArgGlySerLeu 141
                                                                                                                                                                                                                41
                                                                                                                                                                                                      AlaGlnGlnThrArgGlyGluGluGlyCysGlnGluThrSerGlnThrGlyArgAspLys
                                                                                                                                                                                                                                              244 TGCGTCAACGCGTGTGTTGGACCGTTTACCATGGTGCTGGAGGTCTTAGCCGGC
                                                                                                                                                                                                                                                                      SerIleAsnGlyValLeuTrpThrValTyrHisGlyAlaGlyThrArgThrIleAlaSer
                                                                                                                                                                                                                                       AsnGlnValGluGlyGluValGlnIleValSerThrAlaThrGlnThrPheLeuAlaThr
      1 (bases 1 to 651)
Malcolm,B.A., Taremi,S.Shane., Weber,P.C. and Yao,N.
Single-chain recombinant complexes of hepatitis C virus )
protease and NS4A cofactor peptide
Patent: US 6211338-A 99 03-APR-2001;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           PAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                           648
                                                                                                                                                                                                                                                                                                                                                                                                                                                            ValaspPheIleProValGluSerLeuGluThrThrMetArgSer 196
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           linear
                                                                                                                                                                                                                                                                                                                                                                                                                                                                    651
165
17
10
                                                                                                   Length:
Matches:
Conservative:
Mismatches:
Indels:
Gaps:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            DNA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           51 bp 1
6211338.
                                                                      144
                                                                                                                                                          (1-651)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       AR145252 651
Sequence 93 from patent US 62
AR145252
                                              Location/Qualifiers
1..651
                                                             /organism="unknown"
187 c 200 g
                                                                                                                                                          US-09-965-594-18 (1-197) x AR145258
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  AR145252.1 GI:15107119
                                                                                                   2.12e-65
871.50
93.33*
84.62*
85.69*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Unclassified.
1 (bases 1 to 651)
                                                                                                                    Percent Similarity:
Best Local Similarity:
                                                                      120 a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Unknown.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Unknown
                                                                                            Alignment Scores:
                                                                                                                                                                                                        22
                                                                                                                                                                                                                                      42
                                                                                                                                                                                                                                                                    62
                                                                                                                                                                                                                                                                                                                                 102
                                                                                                                                                                                                                                                                                                                                                 364
                                                                                                                                                                                                                                                                                                                                                                 122
                                                                                                                                                                                                                                                                                                                                                                               434
                                                                                                                                 Query Match:
DB:
                                                        source
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  DEFINITION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ACCESSION
VERSION
KEYWORDS
SOURCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ORGANISM
                                                                      BASE COUNT
ORIGIN
                                                                                                     Pred. No.:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           RESULT 13
AR145252
LOCUS
               AUTHORS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          REFERENCE
                                        JOURNAL,
                                               FEATURES
                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Q
                                                                                                                                                                                                                                                                                   đ
                                                                                                                                                                                                                                                                                                               a
                                                                                                                                                                                                                                                                                                                                                                            Q
                                                                                                                                                                                                        ò
                                                                                                                                                                                                                     a
                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                     a
                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                   à
                                                                                                                                                                                                                                                                                                                                ç
                                                                                                                                                                                                                                                                                                                                                a
                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                               à
                                                                                                                                                                         ò
                                                                                                                                                                                        ద
```

```
PAT 08-AUG-2001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 423
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   141
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   483
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   161
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            243
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                303
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              101
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               304 CCAAAGGGGCCAATCACCCAGATGTACACTAATGTGGACCAGGACCTCGTCGGCTGGCAG 363
                                                                                                                                                                                                                                                                                                                                                          64 GGTTCTGTTATTGTTGGTAGAATTATTTTATCTGGTAGTGGTAGTATCACGCCTAC 123
                                                                                                                                                                                                                                                                                                                                                                                                                                            124 TCCCAACAGGGGGGGCGTACTTGGTTGCAAGATCACTAGCCTTACAGGCCGGGACAAG 183
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   AlaGlyHisAlaValGlyIlePheArgAlaAlaValSerThrArgGlyValAlaLysAla 181
                                                                                                                                                                                                                                                                                                                                                                                                            41
                                                                                                                                                                                                                                                                                                                                                                                                                                                                              61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              81
                                                                                                                                                                                                                                                                                                                                             21
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SerileAsnGlyValLeuTrpThrValTyrHisGlyAlaGlyThrArgThrileAlaSer ::::||{|||||||||
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              244 TGCGTCAACGGCGTGTTTGGACCGTTTACCATGGTGCTGGTCAAAGACCTTAGCCGGC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               AlaProGlnGlySerArgSerLeuThrProCysThrCysGlySerSerAspLeuTyrLeu
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  LeuSerProArgProIleSerTyrLeuLysGlySerSerGlyGlyProLeuLeuCysPro
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   484 CTCTCCCCCAGGCCTGTCTCCTACTTGAAGGCTCTTCGGGTGGTCCACTGCTCTGCCCT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ProLysGlyProValThrGlnMetTyrThrAsnValAspLysAspLeuValGlyTrpGln
                                                                                                                                                                                                                                                                                                                                          GlySerValValIleValGlyArgIleAsnLeuSerGlyAsp-----ThrAlaTyr
                                                                                                                                                                                                                                                                                                                                                                                                          22 AlaGlnGlnThrArgGlyGluGluGlyCysGlnGluThrSerGlnThrGlyArgAspLys
                                                                                                                                                                                                                                                                                                                                                                                                                                                                         AsnGlnValGluGlyGluValGlnIleValSerThrAlaThrGlnThrPheLeuAlaThr
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 NS3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1 (bases 1 to 651)
Malcolm, B.A., Tarem1, S.Shane., Weber, P.C. and Yao, N.
Single-chain recombinant complexes of hepatitis C virus
Malcolm, B.A., Taremi, S. Shane., Weber, P.C. and Yao, N. Single-chain recombinant complexes of hepatitis C virus protease and NG4A cofactor peptide patent: US 6211338-A 93 03-APR-2001;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     182 ValAspPheIleProValGluSerLeuGluThrThrMetArgSer 196
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         linear
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   651
115
113
111
                                                                                                                                                                                      Length:
Matches:
Conservative:
Mismatches:
Indels:
Gaps:
                                                                                                                       ų
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       AR145253 651 bp
Sequence 94 from patent US 6211338.
AR145253. GI:15107120
                                                                                                                       145
                                                                                                                                                                                                                                                                                                          US-09-965-594-18 (1-197) x AR145252 (1-651)
                                                                     Location/Qualifiers
                                                                                                  /organism="unknown"
188 c 199 g
                                                                                                                                                                                      2.58e-65
870.50
92.82%
85.13%
85.59%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Unknown.
Unclassified
                                                                                                                                                                                                                      Percent Similarity:
Best Local Similarity:
Query Match:
DB:
                                                                                                                     æ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Unknown.
                                                                                                                                                                       Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   162
                                                                                                                                                                                                                                                                                                                                                                                                                                                                              42
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              184
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              62
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              82
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               102
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 364
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   122
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 424
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 142
```

σ

```
Alignment Scores:
                                                                                                                                                                                                                                                                                                                                        122
                                                                                                                                                                                                                                                                                                                                                                                                                                                           604
                                                                                                                                                                                                     42
                                                                                                                                                                                                                                                                                      304
                                                                                                                                                                                                                                                                                                       102
                                                                                                                                                                                                                                                                                                                                                        424
                                                                                                                                                                                                                                                                       82
                                                                                                                                                                                                                                                                                                                        364
                                                                                          Query Match:
          source
                                                            :.
2
                         COUNT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Search co
 FEATURES
                                   ORIGIN
                           BASE
                                                            Pred
                                                                                                                                                   g
                                                                                                                                                                                g
                                                                                                                                                                                                                  a
                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                     a
                                                                                                                                                                                                                                                                                     ď
                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                        ŏ
                                                                                                                                                                                                                                                                                                                                                        셤
                                                                                                                                                                                                                                                                                                                                                                          à
                                                                                                                                                                                                                                                                                                                                                                                          셤
                                                                                                                                                                                                                                                                                                                                                                                                          ö
                                                                                                                                                                                                                                                                                                                                                                                                                         d
                                                                                                                                                                                                                                                                                                                                                                                                                                           à
                                                                                                                                    à
                                                                                                                                                                    õ
                                                                                                                                                                                                     à
                                                                                                                                                                                                                                                                      à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            08-AUG-2001
                                                                                                                                                                                                                                                                                                                                                        ValThrangHisAlaaspValIleProValArgArgArgGlyAspSerArgGlySerLeu 141
                                                                                                                                                                                                                                                                                                                                                                                                          543
                                                                                                                                                    243
                                                                                                                                                                                                                                                    82 ProLysGlyProValThrGlnMetTyrThrAsnValAspLysAspLeuValGlyTrpGln 101
                                                                                                                                                                                                                                                                                                       363
                                                                                                                                                                                                                                                                                                                        AlaProGlnGlySerArgSerLeuThrProCysThrCysGlySerSerAspLeuTyrLeu 121
                                                                                                                                                                                                                                                                                                                                GICACGAGACAIGCIGACGICAIICCGGIGCGCGGCGGGGGGGACAGIAGGGGGAGCCIG 483
                                                                                                                                                                                                                                                                                                                                                                                          161
                                                                                                                                                                                                                                                                                                                                                                                                                          162 AlaGlyHisAlaValGlyIlePheArgAlaAlaValSerThrArgGlyValAlaLysAla 181
                                                                                                                                                                                                                                                                                                                                                                                                                                   41
                                                                                                                                                                                                                      61
                                                                                                                                                                                                                                                                                                                                                                                          LeuSerProArgProlleSerTyrLeuLysGlySerSerGlyGlyProLeuLeuCysPro
                                                                                                                                                                                                                                                                                                                                                                                                  AsnGlnValGluGlyGluValGlnIleValSerThrAlaThrGlnThrPheLeuAlaThr
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1 (bases 1 to 1998)
Malcolm,B.A., Taremi,S.Shane., Weber,P.C. and Yao,N.
Single-chain recombinant complexes of hepatitis C virus NS3
protease and N84e offactor peptide
Patent: US 6211338-A 106 03-APR-2001,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             PAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                          linear
                                                                           651
166
15
11
13
                                                                           Length:
Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Sequence 106 from patent US 6211338.
ARI45265
protease and NS4A cofactor peptide
Patent: US 6211338-A 94 03-APR-2001;
Location/Qualifiers
                                                                                                                   Gaps:
                                                                                                                                    US-09-965-594-18 (1-197) x AR145253 (1-651)
                                 /organism="unknown"
188 c 199 g
                                           199 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     AR145265.1 GI:15107132
                                                                           2.58e-65
870.50
92.82%
85.13%
85.59%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Unclassified.
                                                                                          Percent Similarity:
Best Local Similarity:
Query Match:
                                            Ø
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Unknown
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Unknown
                                                                    Alignment Scores:
                                                                                                                                                                                                                                                                                                                                       364
                                                                                                                                                                                                                                                                                                                                                                                          142
                                                                                                                                                                                                                      42
                                                                                                                                                                                                                                                                                                       304
                                                                                                                                                                                                                                                                                                                        102
                                                                                                                                                                                                                                                                                                                                                         122
                                                                                                                                                                                                                                                                                                                                                                          124
                                                                                                                                                                                                                                                                                                                                                                                                          184
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   AR145265
LOCUS
DEFINITION
                            source
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ACCESSION
VERSION
KEYWORDS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ORGANISM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              REFERENCE
AUTHORS
                                                                             Pred. No.:
                                           BASE COUNT
        JOURNAL
FEATURES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            RESULT 15
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                JOURNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 TITLE
                                                    ORIGIN
                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                      Ω
                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                     셤
                                                                                                                                                                                                                                                                                                                                                        δ
                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                          õ
                                                                                                                                                      ò
                                                                                                                                                                     셤
                                                                                                                                                                                    ò
                                                                                                                                                                                                  g
                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                      5
                                                                                                                                                                                                                                                                                    õ
                                                                                                                                                                                                                                                                                                                                                                         ద
                                                                                                                                                                                                                                                                                                                                                                                       δ
```

```
ValThrArgHisAlaAspValIleProValArgArgArgGlyAspSerArgGlySerLeu 141
                                                                                                                                                                                                                                                                                                                                                                                              181
                                                                                                                                                                                                                                                                                                                                                                                                             603
                                                                                                                             243
                                                                                                                                                                                               61
                                                                                                                                                                                                                                                                                                                                                                                             AsnGlnValGluGlyGluValGlnIleValSerThrAlaThrGlnThrPheLeuAlaThr
                                                                                                                                                                                                                                                                                                                                                                                                                                              651
                                                                                                                                                                                                                                                                                                                                                                                                                              182 ValAspPheIleProValGluSerLeuGluThrThrMetArgSerPro 197
                                                                                                                                                                                                                                                                                                                                                                                                                                       1998
166
15
12
3
                                                                        Conservative:
Mismatches:
Indels:
                                                        Length:
Matches:
                         425
                                                                                                               US-09-965-594-18 (1-197) x AR145265 (1-1998)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                     completed: August 31, 2003, 00:46:23
He : 2569.57 secs
Location/Qualifiers
1. .1998
/organism="unknown"
a 597 c 567 g
                                                        9.53e-65
870.50
92.35%
84.69%
85.59%
                                                                               Best Local Similarity:
                         409 a
                                                                        Percent Similarity:
```

```
GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.
```

OM protein - nucleic search, using frame_plus_p2n model	
Run on: August 30, 2003, 19:13:57 ; Search time 182.939 Seconds (without alignments) 2906.924 Million cell updates/sec	
Title: US-09-965-594-18 Perfect score: 1017 Sequence: 1 MKKKGSVVIVGRINLSGDTAVAKAVDFIPVESLETTMRSP 197	
Scoring table: BLOSUM62	
Searched: 2552756 seqs, 1349719017 residues	
Total number of hits satisfying chosen parameters: \$105512	
Minimum DB seq length: 0	
Post-processing: Minimum Match 10% Maximum Match 100% Listing first 45 summaries	
Command line parameters: -WODEL-frame+_p2n.model -DEV-xlp -Q-GOR2_L/GSFO_SPROL/VG09965594/runat_29082003_151918_28302/app_query.fasta_1.2872 -Q-GOR2_L/GSFO_SPROL/VG09965594/runat_29082003_151918_28302/app_query.fasta_1.2872 -Q-GORZ_L/GW-Fastap -SUFFIx-rng -MINAATCH+0.1 -LOOPCL-0 -LOOPEXT-0 -UNITS-bits -START-1 -END1 -MATRIX-blosum62 -TRANS-human40.cdi -LIST-45 -DOCALIGW-200 -THR_SCORE-pct -THR_XAX-100 -THR_MTH-0 -ALIGN-15 -NODE-LOCAL -OUTFWT-pto -NORM-ext -HEAPSIZE-500 -MINEN-0 -MINEN-200000000 -USER-USO9955594_GCGN_1 1_1412_€runat_29082003_151918_28302 -NCPU-6 -ICPU-3 -NO_MAAP -LARGEQUERY -NGG_SCORES-0 -WAIT -DSPBLOCK-100 -LONGLOG -DEV_INMOUT-120 -WARN_INMEOUT-30 -THREADS-1 -XGAPOP-10 -XGAPEXT-0.5 -FGAPOP-6 -PEGAPEXT-7 -YGAPOP-10 -YGAPEXT-0.5 -DBLORT-7	N
Database: N_Geneseq_19Jun03:*    SIDSI/gcgdata/geneseqr-embl/NA1980.DAT:*   SIDSI/gcgdata/geneseqr-embl/NA1981.DAT:*   SIDSI/gcgdata/geneseqr-embl/NA1981.DAT:*   SIDSI/gcgdata/geneseqr-embl/NA1981.DAT:*   SIDSI/gcgdata/geneseqr/ember-embl/NA1981.DAT:*   SIDSI/gcgdata/geneseqr-embl/NA1981.DAT:*   SIDSI/gcgdata/geneseqr-embl/NA1981.DAT:*   SIDSI/gcgdata/geneseqr-embl/NA1981.DAT:*   SIDSI/gcgdata/geneseqr-embl/NA1981.DAT:*   SIDSI/gcgdata/geneseqr-embl/NA1981.DAT:*   SIDSI/gcgdata/geneseqr-embl/NA1981.DAT:*   SIDSI/gcgdata/geneseqr/geneseqr-embl/NA1981.DAT:*   SIDSI/gcgdata/geneseqr/geneseqr-embl/NA1981.DAT:*   SIDSI/gcgdata/geneseqr/geneseqr-embl/NA1981.DAT:*   SIDSI/gcgdata/geneseqr/geneseqr-embl/NA1991.DAT:*   SIDSI/gcgdata/geneseqr/geneseqr-embl/NA1991.DAT:*   SIDSI/gcgdata/geneseqr/geneseqr-embl/NA1991.DAT:*   SIDSI/gcgdata/geneseqr/geneseqr-embl/NA1991.DAT:*   SIDSI/gcgdata/geneseqr/geneseqr-embl/NA1991.DAT:*   SIDSI/gcgdata/geneseqr/geneseqr-embl/NA1991.DAT:*   SIDSI/gcgdata/geneseqr/geneseqr-embl/NA1999.DAT:*   SIDSI/gcgdata/geneseqr/geneseqr-embl/NA1999.DAT:*   SIDSI/gcgdata/geneseqr/geneseqr-embl/NA1999.DAT:*   SIDSI/gcgdata/geneseqr/geneseqr-embl/NA1999.DAT:*   SIDSI/gcgdata/geneseqr/geneseqr-embl/NA1999.DAT:*   SIDSI/gcgdata/geneseqr/geneseqr-embl/NA1999.DAT:*   SIDSI/gcgdata/geneseqr/geneseqr-embl/NA2001B.DAT:*   SIDSI/gcgdata/geneseqr-embl/geneseqr-embl/NA2001B.DAT:*   SIDSI/g	

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed,

and is derived by analysis of the total score distribution.

SUMMARIES  Description	TO AT TO SOC	epat	Hepatitis C viru	Hepatitis C viru	Repatitis C viru	Hebatitis C	Hepatit	Hepatitis C v	Chimeric ByDy	DON-RAPA VOH	HCV NS4A-NS3	HCV NS4A-NS3	Anti-viral sy	HCV NS4A-NS3	HCV NS4A-NS	HCV NS4A-NS3	HCV NS4A-NS3	HCV NS4A-NS3	HCV NS4A-NS3	HCV NS4A-NS3	HCV NS4A-NS	HCV NS4A-NS3	HCV NS4A-NS3	HCV NS4A-NS3	HCV NG4A-NG	HCV NS4A-NS3	HCV NS4A-NS3	HCV NOGA-NO	HCV NS3 DNA.	Hepatitis C v	cDNA encoding	AAV39354 Hepatitis C virus ABK87286 Hepatitis C virus	Nucleotide se	Combined ope	Hepatitis C Virus	Combined open r	Composite hepatit	ALIGNMENTS			BP.			usion protease coding sequence #5.	: chronic liver di	utant; mutein; ds.		
ength DB		594						588	734		866	866		866		855	510		651		64.0					648 20		145	933	646	9646	200	6622	300	000		310				DNA; 594		t entry)	NS4A-NS3 f	otease: vir	r cancer;		
% Ouery Match		100.0	86	98	2 9	9 6	93	88	87	98	98	98	9 9	80	96	200	0 00 0 10 0 10	8 6	85	80.0	e e	8 2	85	80 0	0 G	85.0	8	0 00	84	84	800	20.00	8	8	6 0	8	84				standard;		000 (firs	C virus	NS3 pr	lure; li	C virus	
Score	,	10	10	70	no	n or	o	6	892	878	877	877	8 4 4	874	874	8/3	87.1	870	870	870	, a	œ	867	867	866		863	8	œ	828	80.0	858	828	854	2 2	854	854		-	332	4A73332	AAA73332;	9-DEC-5	lepatitis	epatiti	iver fai	Hepatitis Synthetic	1
Result No.		1	(3)	ω <u>~</u>	op L	n vo	7	<b>&amp;</b> (	ع. د	11	12	13	14	16	17	80 5	100	21	22	23	25	26	27	28	30	31	32	9.60	35	36	37	0 F	4	41	4 4	44	45		E-14040	AAA733:	ID A		L TO	==	x	-	OS H	

us-09-965-594-18.rng

```
The present sequence is the coding sequence for a mutated version of a fusion protein created using the Hepatitis C virus (HCV) NS3 and NS4A protease enzymes. These proteins are both essential for the replication of the virus, acting to cleave its replicative proteins from the polyprotein produced from the HCV genome. Inhibitors of the two proteins should be effective as antiviral treatments of HCV infection. This is useful as HCV can lead to chronic liver disease such as cirrhosis, liver failure and liver cancer. The present invention concerns a number of NS3 mutants and NS3-NS4A fusion proteins which can be used to identify inhibitors of this type, as well as enabling structural studies of the protease and protease inhibitor complexes. The protein produced from this sequence contains the alpha-helix0-1 variant.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      41 LysasnGlnValGluGlyGluValGlnIleValSerThrAlaThrGlnThrPheLeuAla 60
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   61 ThrSerIleAsnGlyValLeuTrpThrValTyrHisGlyAlaGlyThrArgThrIleAla 80
                                                                                                                                                                                                                                                                                              Modified hepatitis C virus (HCV) NS3 protease comprising at least 1 substitution of a hydrophobic alpha-helix 0 amino acid to a hydrophilic amino acid, useful for screening inhibitors that may treat hepatitis C
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           21 TyralaGlnGlnThrArgGlyGluGluGlyCysGlnGluThrSerGlnThrGlyArgAsp
                             "NS4A-NS3 fusion protein #5"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Sequence 594 BP; 105 A; 189 C; 153 G; 147 T; 0 other;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       594
0
0
0
0
                                                                                                                                                                                                                    Goldfarb V;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Length:
Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               US-09-965-594-18 (1-197) x AAA73332 (1-594)
                                                                                                                                                                                                                    Weinheimer S, Zhang Y,
                                                                                                                                                                                                                                                                                                                                                                         Claim 26; Fig 15; 66pp; English.
                                                                                                                                                                                    (BRIM ) BRISTOL-MYERS SQUIBB CO
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1,34e-86
1017.00
100.00%
100.00%
                                                                                                                                                       99US-0115271
                                                                                                                         06-JAN-2000; 2000WO-US00345
                               /product-
                                                                                                                                                                                                                                                   WPI; 2000-465976/40.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Percent Similarity:
Best Local Similarity:
                                                                                                                                                                                                                                                                  P-PSDB; AAB15223
                                                            W0200040707-A1
                                                                                                                                                                                                                    Wittekind M,
                                                                                                                                                        38-JAN-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Alignment Scores:
                                                                                          13-JUL-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Query
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ö
g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
```

```
The present sequence is the coding sequence for a mutated version of a fusion protein created using the Hepatitis C virus (HCV) NS3 and NS4A protease enzymes. These proteins are both essential for the replication of the virus, acting to cleave its replicative proteins from the polyprotein produced from the HCV genome. Inhibitors of the two proteins should be effective as antiviral treatments of HCV infection. This is useful as HCV can lead to chronic liver disease such as cirrhosis, liver failure and NS3-NS4A fusion proteins which can be used to identify inhibitors of this type, as well as enabling structural studies of the protease and protease inhibitor complexes. The protein produced from this sequence contains the alpha-helix0-7 variant.
                                          301 CAGGCTCCGCAGGGTTCCCGTTCCCTGACCCCGTGCACCTGCGGTTCCTCCGACCTGTAC 360
                                                                                         LeuLeuSerProArgProIleSerTyrLeuLysGlySerSerGlyGlyProLeuLeuCys 160
                                                                                                       ProAlaGlyHisAlaValGlyIlePheArgAlaAlaValSerThrArgGlyValAlaLys 180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Modified hepatitis C virus (HCV) NS3 protease comprising at least 1 substitution of a hydrophobic alpha-helix 0 amino acid to a hydrophilic amino acid, useful for screening inhibitors that may treat hepatitis C
                                                                                                                                                                     CCGCTCGTCACCCTGTTGCTTCCCGTGCTGTTTCCACCCGTGGTGCTACTAAA
                                                                                                                                                                                                                                                                                                                                                                                                                             Hepatitis, NS3 protease; viral replication; chronic liver disease;
liver failure; liver cancer; mutant; mutein; ds.
                                                                                                                                                                                                               AlaValAspPhelleProValGluSerLeuGluThrThrMetArgSerPro 197
                                                                                                                                                                                                                               Hepatitis C virus NS4A-NS3 fusion protease coding sequence #6.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   /*tag= a
/product= "NS4A-NS3 fusion protein #6"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Goldfarb V;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Zhang Y,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Claim 26; Fig 16; 66pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       (BRIM ) BRISTOL-MYERS SQUIBB CO.
                                                                                                                                                                                                                                                                                                           ВЪ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Weinheimer S,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           06-JAN-2000; 2000WO-US00345.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          99US-0115271
                                                                                                                                                                                                                                                                                                        AAA73333 standard; DNA; 594
                                                                                                                                                                                                                                                                                                                                                                      (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 2000-465976/40.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 WPI; 2000-465976,
P-PSDB; AAB15224.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 WO200040707-A1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          08-JAN-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Wittekind M,
                                                                                                                                                                                                                                                                                                                                                                   19-DEC-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              13-JUL-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Synthetic
                                                                                                                                                   161
                                                                                                                                                                                481
                                                                                                                                                                                                             181
                                                                                                                                                                                                                                           541
                                                                                       141
                                                                                                                    421
                                                                                                                                                                                                                                                                                                                                       AAA73333;
                                                                                                                                                                                                                                                                                        RESULT
                                                                                                                                                                               a
                                                                                                                                                                                                               ò
                                                  셤
                                                                                       Qγ
                                                                                                                    a
                                                                                                                                                   ò
                           οž
```

180

SerProLysGlyProValThrGlnMetTyrThrAsnValAspLysAspLeuValGlyTrp 100

241 81

ö a 101

GlnAlaProGlnGlySerArgSerLeuThrProCysThrCysGlySerSerAspLeuTyr 120

Page

```
WO200040707-A1
                13-JUL-2000
                                                                                                                                                                                                                                                                                                                                41
                                                                                                                                                                                                                                                                                                                                                                                                      101
                                                                                                                                                                                                                                                                                                                                                                              81
                                                                                                                                                                                                                                                                                 à
                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                   a
                                                                                                                                                                                                                                                                                                                                å
                                                                                                                                                                                                                                                                                                                                           q
                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                              à
                                                                                                                                                                                                                                                                                                                                                                                                      à
                                                                                                                                                                                                                                                                                                                                                                                                                 a
                                                                                                                                                                                                                                                                                                                                                                                                                             ò
480
                                                                                                                                                                                            120
                                                                                                                                                                                                                                         LeuLeuSerProArgProlleSerTyrLeuLysGlySerSerGlyGlyProLeuLeuCys 160
                                                                                                                                  180
                                                                                                      61 FACGCTCAGCAGACTCGAGGTGAGCAGGGTTGCCAGAAGACCTCCCACACACGGTCGTGAC 120
                                                                                                40
                                                                                                                       09
                                                                                                                                             80
                                                                                    9
                                                                                                                             ThrSerIleAsnGlyValLeuTrpThrValTyrHisGlyAlaGlyThrArgThrIleAla
                                                                                                                                                                                            GlnAlaProGlnGlySerArgSerLeuThrProCysThrCysGlySerSerAspLeuTyr
                                                                                                                                                                                                                                                LeuValThrArgHisAlaAspValIleProValArgArgArgGlyAspSerArgGlySer
                                                                         TyralaGlnGlnThrargGlyGluGluGlyCysGlnGluThrSerGlnThrGlyArgAsp
                                                                                                                      LysAsnGlnValGluGlyGluValGlnIleValSerThrAlaThrGlnThrPheLeuAla
                                                                                                                                                                                                                                                                                                                                                                        protease; viral replication; chronic liver disease;
liver cancer; mutant; mutein; ds.
                                                                                                                                                                                                                                                                                       Hepatitis C virus NS4A-NS3 fusion protease coding sequence #4.
                                                                                                                                                                                                                                                                                                                                                                                                                             /product= "NS4A-NS3 fusion protein #4"
     other
                      Length:
Matches:
Conservative:
Mismatches:
     0
     ..
H
     G; 147
                                                              (1-594)
                                                                                                                                                                                                                                                                                                                                                                                                            Location/Qualiflers
     152
                                                               US-09-965-594-18 (1-197) x AAA73333
     ပ
                                                                                                                                                                                                                                                                                                                           AAA73331 standard; DNA; 594
     A; 191
                     1.8e-85
1005.00
99.49%
98.48%
                                                                                                                                                                                                                                                                                                                                                  (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                 1..594
/*tag=
     104
                                                                                                                                                                                                                                                                                                                                                                                        Hepatitis C virus.
Synthetic.
                                 Percent Similarity:
Best Local Similarity:
     BP;
                                                                                                                                                                                                                                                                                                                                                                              liver failure;
     Sequence 594
                                                                                                                                                                                                                                                                                                                                                  19-DEC-2000
                 Alignment Scores
                                                                                                                                                                                                                                         141
                                                                                                                                                                                                       301
                                                                                                                                                                                                                               361
                                                                                                                                                                                                                                                                                        181
                                                                                                                                                                                            101
                                                                                                                                                                                                                                                     421
                                                                                                                                                                                                                   121
                                                                                                                                              61
                                              Query Match:
                       ..
Q
                                                                                                                                                                                                                                                                                                               RESULT 3
                                                                                                                                                                                                                                                                                                                    qq
                                                                                     a
                                                                                                ò
                                                                                                           a
                                                                                                                                                                    δý
                                                                                                                                                                               g
                                                                                                                                                                                            ö
                                                                                                                                                                                                     q
                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                ö
                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                     ò
                                                                          ò
                                                                                                                       5
                                                                                                                               Q
                                                                                                                                             δ
                                                                                                                                                         g
```

```
The present sequence is the coding sequence for a mutated version of a fusion protein created using the Hepatitis C virus (HCV) NS3 and NS4A protease enzymes. These proteins are both essential for the replication of the virus, acting to cleave its replicative proteins from the polyprotein produced from the HCV genome. Inhibitors of the two proteins should be effective as antiviral treatments of HCV infection. This is useful as HCV can lead to chronic liver disease such as cirrhosis, liver failure and liver cancer. The present invention concerns a number of NS3 mutants and NS3-NS4A fusion proteins which can be used to identify inhibitors of this type, as well as enabling structural studies of the protease and protease:inhibitor complexes. The protein produced from this sequence contains the alpha-helix0-1 variant.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   GlnAlaProGlnGlySerArgSerLeuThrProCysThrCysGlySerSerAspLeuTyr 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        121 LeuValThrArqHisAlaAspValIleProValArqArqGrqGlyAspSerArqGlySer 140
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  61 TACGCTCAGCAGACTCGAGGTGAGGGTTGCCAAGAAACCTCCCAGACCGGTCGTGAC 120
                                                                                                                                                                                                                                                                                                 Modified hepatitis C virus (HCV) NS3 protease comprising at least 1 substitution of a hydrophobic alpha-helix 0 amino acid to a hydrophilic amino acid, useful for screening inhibitors that may treat hepatitis C
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                21 TyralaGlnGlnThrArgGlyGluGluGlyCysGlnGluThrSerGlnThrGlyArgAsp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         LysasnGlnValGluGlyGluValGlnIleValSerThrAlaThrGlnThrPheLeuAla
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    61 ThrSerIleAsnGlyValLeuTrpThrValTyrHisGlyAlaGlyThrArgThrIleAla
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SerProLysGlyProValThrGlnMetTyrThrAsnValAspLysAspLeuValGlyTrp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 301 CAGGCTCCGCAGGGTTCCCCTGACCCCGTGCACCTGCGGGTTCCTCCGACCTGTAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Sequence 594 BP; 105 A; 187 C; 155 G; 147 T; 0 other;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   594
0 0 0
0 0 0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Conservative:
Mismatches:
Indels:
                                                                                                                                                                 Goldfarb
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Length:
Matches:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               (1-594)
                                                                                                                                                                 Zhang Y,
                                                                                                             8
                                                                                                                                                                                                                                                                                                                                                                                                                                         Claim 26; Fig 14; 66pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               US-09-965-594-18 (1-197) x AAA73331
                                                                                                             (BRIM ) BRISTOL-MYERS SQUIBB
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   3.44e-85
1002.00
98.48%
98.48%
                                                                                                                                                                 ŝ
                                                      99US-0115271
                                                                                                                                                                 Weinheimer
                                                                                                                                                                                                                      2000-465976/40
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Percent Similarity:
Best Local Similarity:
Query Match:
                                                                                                                                                                                                                                                PSDB; AAB15222
                                                                                                                                                                 Wittekind M,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Alignment Scores:
06-JAN-2000;
                                                      08-JAN-1999;
```

```
Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    RESULT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                đ
                                                                                                                                                                                                                                                                               a
                                                                                                                                                                                                                                                                                                                         qq
                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                          셤
                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                 엄
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                  ò
                                                                                                                                        q
                                                                                                                                                              ò
                                                                                                                                                                                q
                                                                                                                                                                                                                                                        õ
                                                                                                                                                                                                                                                                                                      ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      The present sequence is the coding sequence for a mutated version of a fusion protein created using the Hepatitis C virus (HCV) NS3 and NS4A protease enzymes. These proteins are both essential for the replication of the virus, acting to cleave its replicative proteins from the polyprotein produced from the HCV genome. Inhibitors of the two proteins should be effective as antivitaal treatments of HCV infection. This is useful as HCV can lead to chronic liver disease such as cirrhosis, liver fallure and NS3 NS4A fusion proteins which can be used to identify inhibitors of this type, as well as enabling structural studies of the protease and protease inhibitor complexes. The protein produced from this sequence contains the alpha-helix0-7 variant.
                                                          480
                                                                                 180
                                                                                             Modified hepatitis C virus (HCV) NS3 protease comprising at least 1 substitution of a hydrophobic alpha-helix 0 amino acid to a hydrophilic amino acid, useful for screening inhibitors that may treat hepatitis C
ProAlaGlyHisAlaValGlyIlePheArgAlaAlaValSerThrArgGlyValAlaLyS
                                                                                                                                                                                                                                                                                           Hepatitis; NS3 protease: viral replication; chronic liver disease;
liver failure; liver cancer; mutant; mutein; ds.
                                                                                                                             Hepatitis C virus NS4A-NS3 fusion protease coding sequence #7.
                                                                                                                                                                                                                                                                                                                                                                                               'product- "NS4A-NS3 fusion protein #7"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Goldfarb V;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Zhang Y,
                                                                                                                                                                                                                                                                                                                                                              Location/Qualifiers
1..594
/*tag= a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Claim 26; Fig 17; 66pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (BRIM ) BRISTOL-MYERS SQUIBB CO
                                                                                                                                                                                                  BP.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Wittekind M, Weinheimer S,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                    06-JAN-2000; 2000WO-DS00345
                                                                                                                                                                                                AAA73334 standard; DNA; 594
                                                                                                                                                                                                                                               (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            WPI; 2000-465976/40.
P-PSDB; AAB15225.
                                                                                                                                                                                                                                                                                                                             Hepatitis C virus.
Synthetic.
                                                                                                                                                                                                                                                                                                                                                                                                                      WO200040707-A1
                                                                                                                                                                                                                                               19-DEC-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                              13-JUL-2000
                                                                                                                                                                                                                        AAA73334;
                                                                                  161
                                                                                                       481
                                   141
                                                                                                                              181
                                                                                                                                                    541
                                                          121
                                                                                                                                                                            RESULT 4
                                                                                                                                                                                         AAA73334
                                                                                                                                                                                                   셤
                                                          8
                                                                                  õ
                                                                                                                            ŏ
```

Sequence 594 BP; 105 A; 192 C; 151 G; 146 T; 0 other;

```
LeuValThrArgHisAlaAspValIleProValArgArgArgGlyAspSerArgGlySer 140
                                                                                                                                                                                                                                                                                                                                                                                 LeuLeuSerProArgProIleSerTyrLeuLysGlySerSerGlyGlyProLeuLeuCys 160
                                                                                                                                                                                                                                                                                                                                                                                             100
                                                                                                                                                                                                                                                                                                             180
                                                                                                                                                                                                                                      240
                                                                                                                                                                                                                                                                      GlnAlaProGlnGlySerArgSerLeuThrProCysThrCysGlySerSerAspLeuTyr 120
                                                                                                                                                                            9
                                                                                                                                   9
                                                                                                                                                                                                                   80
                                                                                            20
                                                                                                               9
                                                                                                                                                                         1 ATGAAAAAAAAGGATCCGTTGTTATCGTCGGCCGTATCAACCTGTCCGGTGACACGGCT
                                                                                                                                  TyralaglnGlnThrArgGlyGluGluGlyCysGlnGluThrSerGlnThrGlyArgAsp
                                                                                                                                                                                                                                                           SerProLysGlyProValThrGlnMetTyrThrAsnValAspLysAspLeuValGlyTrp
                                                                                          MetLysLysLysGlySerValValIleValGlyArglleAsnLeuSerGlyAspThrAla
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Hepatitis; NS3 protease; viral replication; chronic liver disease;
liver failure; liver cancer; mutant; muteln; ds.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                   AlavalAspPheIleProValGluSerLeuGluThrThrMetArgSerPro 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Hepatitis C virus NS4A-NS3 fusion protease coding sequence #3.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                             /*tag= a
/product= "NS4A-NS3 fusion protein #3"
594
22
20
00
00
                    Conservative:
Mismatches:
Indels:
           Matches:
 Length:
                                                     Gaps:
                                                                       US-09-965-594-18 (1-197) x AAA73334 (1-594)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             AAA73330 standard; DNA; 594 BP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (first entry)
          995.00
98.98%
97.97%
97.84%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Hepatitis C virus.
Synthetic.
                 Percent Similarity:
Best Local Similarity:
Query Match:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            WO200040707-A1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     19-DEC-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      541
                                                                                                                                                                                                                                                                                                  101
                                                                                                                                                                                                                                                                                                                                          121
                                                                                                                                                                                                                                                                                                                                                              361
                                                                                                                                                                                                                                                                                                                                                                                   141
                                                                                                                                                                                                                                                                                                                                                                                                                          161
                                                                                                                                                                                                                                                                                                                                                                                                                                                                   181
                                                                                                                                                                            41
                                                                                                                                                                                                                                                           81
                                                                                                                                                                                                                                                                                                                       301
                                                                                                                                                                                                                                                                                                                                                                                                                                               481
   ..
9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Key
```

```
LeuLeuSerProArgProlleSerTyrLeuLysGlySerSerGlyGlyProLeuLeuCys
                                                                                                                                                                                                       (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Weinheimer
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             2000-465976/40.
                                                                                                                                                                                                                                                                                   Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        P-PSDB; AAB15220
                                                                                                                                                                                                                                                                                                                                                                          W0200040707-A1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Wittekind M,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Alignment Scores:
Pred. No.:
Score:
                                                                                                                                                                                                                                                                                                                                                                                                                                            08-JAN-1999;
                                                                                                                                                                                                       19-DEC-2000
                                                                                                                                                                                                                                                                                                                                                                                                 13-JUL-2000
                                                                                                                                                                                                                                                                                               Synthetic
                                                                                                                                                                                AAA73329;
 141
                       421
                                                                  481
                                                                                                                                                AAA73329
                                                                                                                                    RESULT
                                                                                                                                                                               õ
                                                                 g
                                                                                         ò
                                                                                                                                                                                                                        The present sequence is the coding sequence for a mutated version of a fusion protein created using the Hepatitis C virus (HCV) NS3 and NS4A protease enzymes. These proteins are both essential for the replication of the virus, acting to cleave its replicative proteins from the PCV genome. Inhibitors of the two proteins should be effective as antiviral treatments of HCV infection. This is should be effective as antiviral treatments of HCV infection. This is useful as HCV can lead to chronic liver disease such as cirrhosis, liver failure and liver cancer. The present invention concerns a number of NS3 inhibitors of this type, as well as enabling structural studies of the protease and protease inhibitor complexes. The protein produced from this sequence contains the alpha-helix0-1 variant.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 81 SerProLysGlyProValThrGlnMetTyrThrAsnValAspLysAspLeuValGlyTrp 100
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         101 GlnAlaProGlnGlySerArgSerLeuThrProCysThrCysGlySerSerAspLeuTyr 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              61 TACGCTCAGCAGACTCGAGGTGAGGAGGGTTGCCAAGAAACCTCCCAGACCGGTCGTGAC 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        40
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      09
                                                                                                                                               Modified hepatitis C virus (HCV) NS3 protease comprising at least 1 substitution of a hydrophobic alpha-helix 0 amino acid to a hydrophilic amino acid, useful for screening inhibitors that may treat hepatitis C
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           21 TyrAlaGlnGlnThrArgGlyGluGluGlyCysGlnGluThrSerGlnThrGlyArgAsp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          61 ThrSerIleAsnGlyValLeuTrpThrValTyrHisGlyAlaGlyThrArgThrIleAla
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      181 ACCTGCATCAACGGTGTTTGCTGGACCGTTTACCACGGTGCTGGTACCGTACCATCGCT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     LysAsnGlnValGluGlyGluValGlnIleValSerThrAlaThrGlnThrPheLeuAla
                                                                                                                                                                                                                                                                                                                                                                          Seguence 594 BP; 103 A; 186 C; 156 G; 149 T; 0 other;
                                                                                                                                                                                                                                                                                                                                                                                                         594
191
0
0
0
0
                                                                                                                                                                                                                                                                                                                                                                                                         Length:
Matches:
Conservative:
Mismatches:
Indels:
                                                                                         Goldfarb
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Gaps:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        (1-594)
                                                                                         Zhang Y,
                                                                                                                                                                                                       Claim 26; Fig 13; 66pp; English.
                                                                 (BRIM ) BRISTOL-MYERS SQUIBB CO.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       US-09-965-594-18 (1-197) x AAA73330
                                                                                                                                                                                                                                                                                                                                                                                                       1.36e-83
985.00
96.95%
96.95%
96.85%
                                              99US-0115271.
                                                                                        Wittekind M, Weinheimer S,
                       06-JAN-2000; 2000WO-US00345
                                                                                                               WPI; 2000-465976/40.
                                                                                                                                                                                                                                                                                                                                                                                                                                 Percent Similarity:
Best Local Similarity:
                                                                                                                           P-PSDB; AAB15221
                                                                                                                                                                                                                                                                                                                                                                                                 Alignment Scores:
                                              08-JAN-1999;
13-JUL-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     41
                                                                                                                                                                                                                                                                                                                                                                                                                                                      Ouery Match:
DB:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           용
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
 qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 å
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  d
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ô
```

```
The present sequence is the coding sequence for a mutated version of a fusion protein created using the Hepatitis C virus (HCV) NS3 and NS4A protease enzymes. These proteins are both essential for the replication of the virus, acting to cleave its replicative proteins from the polyprotein produced from the HCV genome. Inhibitors of the two proteins should be effective as antiviral treatments of HCV infection. This is useful as HCV can lead to chronic liver disease such as cirrhosis, liver failure and liver cancer. The present invention concerns a number of NS3 mutants and NS3-NS4A fusion proteins which can be used to identify inhibitors of this type, as well as enabling structural studies of the protease and protease:inhibitor complexes. The protein produced from this
                                  480
                                                                                                180
                                                                                                                                                              540
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Modified hepatitis C virus (HCV) NS3 protease comprising at least 1 substitution of a hydrophobic alpha-helix 0 amino acid to a hydrophilic amino acid, useful for screening inhibitors that may treat hepatitis C
161 ProAlaGlyHisAlaValGlyIlePheArgAlaAlaValSerThrArgGlyValAlaLys
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Hepatitis; NS3 protease; viral replication; chronic liver disease;
liver failure; liver cancer; mutant; mutein; ds.
                                                                                                                                                                                                                     Hepatitis C virus NS4A-NS3 fusion protease coding sequence #2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  /*tag= a
/product= "NS4A-NS3 fusion protein #2"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Sequence 588 BP; 103 A; 180 C; 156 G; 149 T; 0 other;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  588
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Goldfarb
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    sequence contains the alpha-helix0-1 variant.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Zhang Y,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Socation/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          (BRIM ) BRISTOL-MYERS SQUIBB CO.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Claim 26; Fig 12; 66pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                            HP.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ŝ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           06-JAN-2000; 2000WO-US00345.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            99US-0115271
                                                                                                                                                                                                                                                                                                                                                                                                                            AAA73329 standard; DNA; 588
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  2.1e-80
951.00
```

```
LeuValThrArgHisAlaAspValIleProValArgArgArgGlyAspSerArgGlySer 140
                                                                                                                                                                                                                                                                                                                           LeuLeuSerProArgPro1leScrTyrLeuLysGlySerSerGlyGlyProLeuLeuCys 160
                                                                                                   SerProLysGlyProValThrGlnMetTyrThrAsnValAspLysAspLeuValGlyTrp 100
                                                                                                                                                                                                                         .01 GlnAlaProGlnGlySerArgSerLeuThrProCysThrCysGlySerSerAspLeuTyr 120
                                                                                                                                                                                                                                                             CIGCIGICCCCGCGTCCGAICTCCTACCTGAAGGIICCTCCGGTGGTCCGCTGCTGTGC 474
                                                                                                                                                                          MetLysLysLysGlySerValValIleValGlyArgIleAsnLeuSerGlyAspThrAla
                                                                           LysAsnGlnValGluGlyGluValGlnIleValSerThrAlaThrGlnThrPheLeuAla
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Hepatitis: NS3 protease; viral replication; chronic liver disease;
liver failure; liver cancer; mutant; mutein; ds.
                                                                                                                                                                                                                                                                                                                                                                                                    181 AlaValAspPheIleProValGluSerLeuGluThrThrMetArgSerPro 197
                                                                                                                                                                                                                                                                                                                                                                                                              /*tag= a
/product= "NS4A-NS3 fusion protein #8"
 4641
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Hepatitis C virus NS4A-NS3 fusion protease coding
 Conservative:
          Mismatches:
                   Indels:
Gaps:
                                               (1-588)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Location/Qualifiers
1..594
/*tag= a
                                                                                                                                                                                                                                                                                                                                                                                                                                                          ВР
                                                US-09-965-594-18 (1-197) x AAA73329
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           06-JAN-2000; 2000WO-US00345
                                                                                                                                                                                                                                                                                                                                                                                                                                                         AAA73335 standard; DNA; 594
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              (first entry)
95.43%
94.92%
93.51%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Hepatitis C virus.
Synthetic.
 Percent Similarity:
Best Local Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       WO200040707-A1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              19-DEC-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         13-JUL-2000
                                                                                                                                                                                                               81
                                                                                                                                                                                                                                                                                                         355
                                                                                                                                                                                                                                                                                                                            141
                                                                                                                                                                                                                                                                                      121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                            AAA73335;
                  Query Match:
DB:
                                                                                                                                                                                                                                                                                                                                                                                                                                                RESULT 7
                                                                                                                                                                                                                                                                                                                                                                             a
                                                                                                                                                                                                                                                                                                                                                                                                   ö
                                                                                                                                                                                                                                                                                                                                                                                                                  a
                                                                                                                       a
                                                                                                                                         õ
                                                                                                                                                     g
                                                                                                                                                                           ç
                                                                                                                                                                                         g
                                                                                                                                                                                                          à
                                                                                                                                                                                                                               Q
                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                 a
                                                                                                                                                                                                                                                                                     õ
                                                                                                                                                                                                                                                                                                    셤
                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                           셤
                                                                                                                                                                                                                                                                                                                                                             à
                                                                                                    ç
```

```
The present sequence is the coding sequence for a mutated version of a fusion protein created using the Hepatitis C virus (HCV) NS3 and NS4A protease enzymes. These proteins are both essential for the replication of the virus, acting to cleave its replicative proteins from the polyprotein produced from the HCV genome. Inhibitors of the two proteins should be effective as antiviral treatments of HCV infection. This is useful as HCV can lead to chronic liver disease such as cirrhosis, liver failure and liver cancer. The present invention concerns a number of NS3 mutants and NS3 NS4A fusion proteins which can be used to identify inhibitors of this type, as well as enabling structural studies of the protease and protease:inhibitor complexes. The protein produced from this
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               140
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  LeuleuSerProArgProlleSerTyrLeuLysGlySerSerGlyGlyProLeuLeuCys 160
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 61 TACGCTCAGCAGACTCGAGGTCTGCTGGTTGCATCATCACCTCCCTGACCGGTCGTGAC 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   80
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         9
                                                                                                                                                          Modified hepatitis C virus (HCV) NS3 protease comprising at least 1 substitution of a hydrophobic alpha-helix 0 amino acid to a hydrophilic amino acid, useful for screening inhibitors that may treat hepatitis C
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ThrSerIleAsnGlyValLeuTrpThrValTyrHisGlyAlaGlyThrArgThrIleAla
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          MetLysLysLysGjySerValValIleValGlyArgIleAsnLeuSerGlyAspThrAla
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1 ATGAAAAAAAAAGGATCCGTTGTTATCGTCGGCCGTATCAACCTGTCCGGTGACACCGCT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        LysasnGlnValGluGlyGluValGlnIleValSerThrAlaThrGlnThrPheLeuAla
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ACCIGCATCAACGGIGTIIGCIGGACCGITIACCACGGIGCIGGIACCGGIACCAICGCI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   GlnAlaProGlnGlySerArgSerLeuThrProCysThrCysGlySerSerAspLeuTyr
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  LeuValThrArgHisAlaAspValIleProValArgArgArgGlyAspSerArgGlySer
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         21 TyralaGlnGlnThrArqGlyGluGluGlyCysGlnGluThrSerGlnThrGlyArqAsp
                                                                                                                                                                                                                                                                                                                                                                                                                                                 sequence contains the alpha-helix0 wild-type sequence.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           594
186
0
0
0
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Sequence 594 BP; 98 A; 189 C; 153 G; 154 T; 0 other;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Length:
Matches:
Conservative:
Mismatches:
Indels:
                                                                              Goldfarb
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          US-09-965-594-18 (1-197) x AAA73335 (1-594)
                                                                               Zhang Y,
                                                                                                                                                                                                                                        Disclosure; Fig 18; 66pp; English.
                                               8
                                             (BRIM ) BRISTOL-MYERS SQUIBB
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        6.27e-80
946.00
94.42%
94.42%
93.02%
                                                                               Weinheimer S,
               990S-0115271
                                                                                                              2000-465976/40
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Percent Similarity:
Best Local Similarity:
                                                                                                                             P-PSDB; AAB15226
                                                                               Wittekind M,
               08-JAN-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        41
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  81
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  241
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   101
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                361
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  141
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 421
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Query Match:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ద
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   엄
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ô
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
```

^

100

294

234

114

50 54 9

174 80 120

354

140

414

160

474

180

RESULT 셤

a à

```
21 TyrAlaGlnGlnThrArgGlyGluGluGluGlyCysGlnGluThrSerGlnThrGlyArgAsp 40
                                                                                                             101 GlnAlaProGlnGlySerArgSerLeuThrProCysThrCysGlySerSerAspLeuTyr
                                                                                                                                                                                                                                                                                                                                                            CCGCCCCCCAGGGTTCCCCGTCCCTGACCCCGTCCACCTGCGGTTCCTCCGACCTGTAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             161 ProAlaGlyHisAlaValGlyIlePheArgAlaAlaValSerThrArgGlyValAlaLys
                                                   SerProLysGlyProValThrGlnMetTyrThrAsnValAspLysAspLeuValGlyTrp
                                1 MetLysLysLysGlySerValValIleValGlyArgIleAsnLeuSerGlyAspThrAla
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Nucleic acid construct encoding chimeric Hepatitis C Virus (HCV)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       181 AlaValAspPheIleProValGluSerLeuGluThrThrMetArgSerPro 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ds
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Pestivirus; Npro; protease; NS3; screening;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Chimeric - Bovine viral diarrhea virus.
Chimeric - Hepatitis C virus.
   US-09-965-594-18 (1-197) x AAA73328 (1-588)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Chimeric BVDV/HCV NS3-wt sequence.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   BP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ABA95615 standard; DNA; 12734
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                99US-0344456
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Lai VCH, Lau JYN;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            (SCHE ) SCHERING CORP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         WPI; 2002-121103/16.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                25-JUN-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               25-JUN-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   US6326137-B1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             21-MAR-2002
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  04 - DEC - 2001
                                                                                                                                                                                                                                                                                81
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  141
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               415
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      535
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Hong 2,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     RESULT 9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ABA9561
                                                                                             à
                                                                                                                        셤
                                                                                                                                                         ò
                                                                                                                                                                                      a
                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                             a
                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                      ö
                                                                                                                                                                                                                                                                                                                                                                                                                                a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Š
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         윱
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    The present sequence is the coding sequence for a fusion protein created using the Hepatitis C virus (HCV) NG3 and NG4A protease enzymes. These proteins are both essential for the replication of the virus, acting to cleave its replicative proteins from the polyprotein produced from the HCV genome. Inhibitors of the two proteins should be effective as antiviral treatments of HCV infection. This is useful as HCV can lead to chronic liver disease such as cirrhosis, liver failure and liver cancer. The present invention concerns a number of NG3 mutants and NG3-NG4A failure proteins which can be used to identify inhibitors of this type, as
                 Modified hepatitis C virus (HCV) NS3 protease comprising at least 1 substitution of a hydrophobic alpha-helix 0 amino acid to a hydrophilic amino acid, useful for screening inhibitors that may treat hepatitis C
ProAlaGlyHisAlaValGlyIlePheArgAlaAlaValSerThrArgGlyValAlaLys
                                                                                                                                                                                                                                                                             depatitis; NS3 protease; viral replication; chronic liver disease;
                                                                             Hepatitis C virus NS4A-NS3 fusion protease coding sequence #1.
                                                             181 AlaValAspPheIleProValGluSerLeuGluThrThrMetArgSerPro
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               well as enabling structural studies of the protease and
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           588
182
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Sequence 588 BP; 97 A; 183 C; 153 G; 155 T; 0 other;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1557
                                                                                                                                                                                                                                                                                                                                                                                                      /*tag= a
/product= "NS3-NS4A fusion protein"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Zhang Y, Goldfarb V;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Conservative:
Mismatches:
Indels:
Gaps:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Length:
Matches:
                                                                                                                                                                                                                                                                                                                                                                         Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Disclosure; Fig 10; 66pp; English.
                                                                                                                                                                                                                                                                                             liver failure; liver cancer; ds.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (BRIM ) BRISTOL-MYERS SQUIBB CO
                                                                                                                                                         BP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               protease:inhibitor complexes
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       9.68e-77
912.00
92.89%
92.39%
89.68%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Wittekind M, Weinheimer S,
                                                                                                                                                       AAA73328 standard; DNA; 588
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           06-JAN-2000; 2000WO-US00345
                                                                                                                                                                                                                   (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 WPI; 2000-465976/40.
P-PSDB; AAB15212.
                                                                                                                                                                                                                                                                                                                          Hepatitis C virus. Synthetic.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Percent Similarity:
Best Local Similarity:
Query Match:
DB:
                                                                                                                                                                                                                                                                                                                                                                                                                                                 WO200040707-A1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          08-JAN-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Alignment Scores:
Pred. No.:
                                                                                                                                                                                                                   19-DEC-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                13-JUL-2000
 161
                                                                                                                                                                                    AAA73328;
```

us-09-965-594-18.rng

(first entry)

07-SEP-1999

AAX80355;

```
pestivirus genome where the Npro protease gene is replaced with NS3 protease gene, useful for in vivo screening of compounds which inhibit HCV infection
                                   Example 2; Columns 17-28; 20pp; English.
```

The present invention relates to a nucleic acid construct encoding a chimeric Hepatitis C virus (HCV)-peativirus genome. The construct comprises a pestivirus genome where a Npro pestivirus protease gene is replaced with a genome encoding a functional HCV NS3 protease. Furthermore, each junction site recognised by the Npro protease is replaced with a function site recognised by the HCV NS3 protease. The construct is useful for screening compounds that inhibit HCV in vivo by inhibiting HCV protease, where screening may be in cell culture or in an animal model. The present sequence is a chimeric clone of BVDV (bovine viral diarrhea The present sequence is a chimeric clone of BVDV (bovine viral diarrhe virus)/ECV NS3-wt, which was used to illustrate the present invention.

Sequence 12734 BP; 4032 A; 2604 C; 3295 G; 2803 T; 0 other;

```
772
                                                                                                                                                                                                                                                                                                            413 GGTAGTGTTGTTGTTGGTAGAATTGTTTTATCTGGTAGTGGTAGTATCACGGCGTAC 472
                                                                                                                                        592
                                                                                                                                                                                                                     652
                                                                                                                                                                                                                                                                          102 AlaProGlnGlySerArgSerLeuThrProCysThrCysGlySerSerAspLeuTyrLeu 121
                                                                                                                                                                                                                                                                                                                                                                                     162 AlaGlyHisAlaValGlyIlePheArgAlaAlaValSerThrArgGlyValAlaLysAla 181
                                                                                                                                                                                                                                                                                                                                                                                              ProLysGlyProValThrGlnMetTyrThrAsnValAspLysAspLeuValGlyTrpGln 101
                                                                                                                                                                                                   81
                                                                                        GlyServalvall1evalGlyarg1leAsnLeuSerGlyAsp------ThrAlaTyr 21
                                                                                                                            41
                                                                                                                                                                42 AsnGlnValGluGlyGluValGlnIleValSerThrAlaThrGlnThrPheLeuAlaThr 61
                                                                                                                                                                                                  22 AlaGlnGlnThrArgGlyGluGluGlyCysGlnGluThrSerGlnThrGlyArgAspLys
                                                                                                                                                                                                                                                                                                                                                                                                                           ValAspPheIleProValGluSerLeuGluThrThrMetArgSer 196
                                                                                                                                                                                                                                                                                                                                                                                                                                             997
                                                                                                                                                                                                                                                                                                                                                                                                                                   Length:
Matches:
Conservative:
Mismatches:
Indels:
                                                                       US-09-965-594-18 (1-197) x ABA95615 (1-12734)
     3.06e-73
892.50
92.82%
90.77%
87.76%
                                   Similarity:
                          Percent Similarity:
Alignment Scores:
                                                                                         S
                                                                                                                                                                                                                                     82
                                                                                                                                                                                                                                                                                                                                                                                                                                            953
                                           Query Match:
DB:
                                                                                                                                                                                                                                                                                                                                                                                                                                                               RESULT 10
                                                                                                         a
                                                                                                                                          a
                                                                                                                                                                                                                                                     8
                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                        ŏ
                                                                                                                                                                                                                                                                                                                            q
                                                                                                                                                                                                                                                                                                                                                 ö
                                                                                                                                                                                                                                                                                                                                                             Q
                                                                                                                                                                                                                                                                                                                                                                                    à
                                                                                                                                                                                                                                                                                                                                                                                                      8
                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                        ò
                                                                                                                           à
                                                                                                                                                               ò
                                                                                                                                                                             g
                                                                                                                                                                                                  ò
                                                                                                                                                                                                                   a
                                                                                                                                                                                                                                      õ
```

```
The present invention describes a covalent hepatitis C virus (MCV) NS4A NS3 complex comprising a central hydrophobic domain of native HCV NS4A peptide, a linker, and an HCV NS3 estine protease domain, where the hydrophobic domain of native HCV NS4A peptide is tethered by the linker to the amino terminus of the HCV NS4A peptide is tethered by the linker to the amino terminus of the HCV NS3A protease domain. The present Sequence encodes an example of the above complex. The covalent NS4A-NS3 complexes are useful for structural determination and determination of mode of binding of HCV inhibitors by NMR spectroscopy. They can also be used for detecting inhibitors of the protease activity, the helicase activity and the ATPase activity of NS3. The covalent covalent complexes are more soluble, stable and active than the non-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          64 GGTTCTGTTGTTGTTGGTAGAATTATTTTATCTGGTAGTGGTAGTATCACGGCCTAC 123
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        22 AlaGlnGlnThrArgGlyGluGluGlyCysGlnGluThrSerGlnThrGlyArgAspLys 41
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              GlyserValValIleValGlyArgIleAsnLeuSerGlyAsp-----ThrAlaTyr
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         42 AsnGlnValGluGlyGluValGlnIleValSerThrAlaThrGlnThrPheLeuAlaThr
                                                                                      HCV; hepatitis C virus; single chain recombinant complex; linker;
NS3 protease; NS4A cofactor; NS4A-NS3 complex; serine protease;
hydrophobic domain; covalent complex; detection; inhibitor; ss.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1998
167
16
10
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Sequence 1998 BP; 411 A; 595 C; 569 G; 423 T; 0 other;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Length:
Matches:
Conservative:
Mismatches:
Indels:
                                                          CDNA SEQ ID NO:105
                                                                                                                                                                                                                                                                                                                                                                   ż
                                                                                                                                                                                                                                                                                                                                                                                                                                                             Disclosure; Page 166-169; 211pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                               New hepatitis C virus covalent complexes
                                                                                                                                                                                                                                                                                                                                                                     Yao
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               US-09-965-594-18 (1-197) x AAX80355 (1-1998)
                                                                                                                                                                                                                                                                                                                                                                   PC,
                                                                                                                                                                                                                                                                                                                                                                     Weber
                                                            complex encoding
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    3.27e-73
881.50
93.37%
85.20%
86.68%
                                                                                                                                                                                                                                                                                       98US-0094331
97US-0067315
                                                                                                                                                                                                                                                           98WO-US2452B
                                                                                                                                                                                                                                                                                                                                                                   Taremi SS,
                                                                                                                                                                                                                                                                                                                                      (SCHE ) SCHERING CORP.
                                                                                                                                                                                                                                                                                                                                                                                                  WPI; 1999-385385/32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Percent Similarity:
Best Local Similarity:
                                                                                                                                                   Hepatitis C virus
                                                                                                                                                                                                                                                                                       28-JUL-1998;
28-NOV-1997;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Alignment Scores:
                                                            HCV NS4A-NS3
                                                                                                                                                                                               W09928482-A2
                                                                                                                                                                                                                                                           24 - NOV - 1998;
                                                                                                                                                                                                                                                                                                                                                                     Malcolm BA,
                                                                                                                                                                                                                            10-JUN-1999
                                                                                                                                                                  Synthetic
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Query Match:
DB:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Score:
ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ç
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       윱
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   a
```

ογ

standard; cDNA; 1998

AAX80355

g

```
Sequence 1998
                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                             q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ç
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             쉽
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                        ö
                                                                                                                                             ò
                                                                                                                                                                    g
                                                                                                                                                                                           ò
                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                     603
                                                                                                           NS4A-NS3 complex comprising a central hydrophobic domain of native RCV NS4A-NS3 complex comprising a central hydrophobic domain, where the hydrophobic domain of native HCV NS4 serine protease domain, where the hydrophobic domain of native HCV NS4 peptide is tethered by the linker to the amino terminus of the HCV NS3 protease domain. The present sequence encodes an example of the above complex. The covalent NS4A-NS3 complexes are useful for structural determination and determination of mode of binding of HCV inhibitors by NMR spectroscopy. They can also be used for detecting inhibitors of the protease activity and the Arpsase activity of NS3. The covalent covalent covalent protease speptide complexes previously available.
            121
                                                                      364 GCGCCCCCCGGGGCGCGTTCCTTGACACCATGCACCTGTGGCAGCTCAGACCTTTACTTG 423
                                                                                              141
                                                                                                                                             LeuSerProArgProlleSerTyrLeuLysGlySerSerGlyGlyProLeuLeuCysPro 161
                                                                                                                                                                                           162 AlaGiyHisAlaValGlyIlePheArgAlaAlaValSerThrArgGlyValAlaLySAla 181
ProLysGlyProValThrGlnMetTyrThrAsnValAspLysAspLeuValGlyTrpGln 101
                                               ValThrArgHisAlaAspValIleProValArgArgArgGlyAspSerArgGlySerLeu
                                                                                                                                                          present invention describes a covalent hepatitis C virus (MCV)
                                                                                                                                                                                                                                                                                                                                                                                                              HCV; hepatitis C virus; single chain recombinant complex; linker;
NS3 protease; NS4A cofactor; NS4A-NS3 complex; serine protease;
hydrophobic domain; covalent complex; detection; inhibitor; ss.
                                                                                                                                                                                                                                             ValAspPheIleProValGluSerLeuGluThrThrMetArgSerPro 197
                                                                                                                                                                                                                                                       ACV NS4A-NS3 complex encoding cDNA SEQ ID NO:109.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Ż
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Disclosure; Page 179-182; 211pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 New hepatitis C virus covalent complexes
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Š,
                                                                                                                                                                                                                                                                                                                   AAX80359 standard; cDNA; 1998
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        98US-0094331.
97US-0067315.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  98WO-US24528
                                                                                                                                                                                                                                                                                                                                                                  (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Taremi SS,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            (SCHE ) SCHERING CORP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         WPI; 1999-385385/32.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                Hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  WO9928482-A2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 24-NOV-1998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        28-JUL-1998;
28-NOV-1997;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Malcolm BA,
                                                                                                                                                                                                                                                                                                                                                                 07-SEP-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          10-JUN-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Synthetic.
                                                102
                                                                                                                     424
                                                                                                                                             142
                                                                                                                                                                                                                                             182
                                                                                                                                                                                                                                                                    604
                                                                                                                                                                                                                                                                                                                                          AAX80359;
                                                                                                                                                                                                                                                                                                    AAX80359
                                                                                                                                                                                                                                                                                            RESULT
                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                            g G
                                                                                                                                                                                      à
                                                                                                                                                                                                              a
                                                                                                                                                                                                                                                               용
                   셤
                                               à
                                                                   ä
                                                                                             ç
                                                                                                                                                                                                                                            ò
```

```
363
                                                                                                                                                                                                                                                                                                                                                                                                                   121
                                                                                                                                                                                                                                                                                                                                                                                                                                            423
                                                                                                                                                                                                                                                                                                                                                                                                                                                                      141
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                483
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       161
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    603
                                                                                                                                                                                                                                                                                                                                        303
                                                                                                                                                                                                                                                                                                                                                                  101
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           162 AlaGlyHisAlaValGlyIlePheArgAlaAlaValSerThrArgGlyValAlaLySAla 181
                                                                                                                                                                                                                                   183
                                                                                                                                                                    22 AlaGlnGlnThrArgGlyGluGluGlyCysGlnGluThrSerGlnThrGlyArgAspLys 41
                                                                                                                                                                                                                                                                                                               83
                                                                                                                                                                                                                                                           122 ValThrargHisAlaAspValIleProValArgArgArgGlyAspSerArgGlySerLeu
                                                                                                                                                                                                                                                                                                                                        244 IGCGICAACGCCGIGIGIIGGACCGIIIACCAIGGIGCIGCIGCICAAAGACCIIAGCCGGC
                                                                                                                                                                                                                                                                                                                                                                  82 ProLysGlyProValThrGlnMetTyrThrAsnValAspLysAspLeuValGlyTrpGln
                                                                                                                                                                                                                                                                                                                                                                               102 AlaproGlnGlySerArgSerLeuThrProCysThrCysGlySerSerAspLeuTyrLeu
                                                                                                                                                                                                                                                                                                                                                                                                                                  SerileAsnGlyvalLeuTrpThrValTyrHisGlyalaGlyThrArgThrileAlaSer
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                424 GTCACGAGACATGCTGACGTCATTCCGGTGCGCCGGCGGGGGGCGACAGTAGGGGGGGCCTG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       LeuSerProArgProlleSerTyrLeuLysGlySerSerGlyGlyProLeuLeuCysPro
                                                                                                                                                        GlySerValValIleValGlyArgIleAsnLeuSerGlyAsp------ThrAlaTyr
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            HCV; hepatitis C virus; single chain recombinant complex; linker; NS3 protease; NS4A cofactor; NS4A-NS3 complex; serine protease; hydrophobic domain; covalent complex; detection; inhibitor; ss.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             182 ValAspPheIleProValGluSerLeuGluThrThrMetArgSerPro 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         G; 423 T; 0 other;
                                   1998
166
17
10
3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     HCV NS4A-NS3 complex encoding cDNA SEQ ID NO:103.
                                     Length:
Matches:
Conservative:
Mismatches:
Indels:
                                                                                                     Gaps:
                                                                                                                            (1-1998)
569
ပ်
                                                                                                                              US-09-965-594-18 (1-197) x AAX80359
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          AAX80353 standard; cDNA; 1998
BP; 411 A; 595
                                  6.26e-73
878.50
93.37%
84.69%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Hepatitis C virus
Synthetic.
                                                                           Best Local Similarity:
                                                               Percent Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           07-SEP-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        WO9928482-A2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  10-JUN-1999
                          Alignment Scores
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     604
                                                                                                                                                                                                                                                                                     184
                                                                                                                                                                                                                                                                                                               62
                                                                                                                                                                                                                                                                                                                                                                                          304
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       142
                                                                                         Query Match:
                                         ..
8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                RESULT 12
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            AAX80353
```

```
22 AlaGlnGlnThrArgGlyGluGluGlyCysGlnGluThrSerGlnThrGlyArgAspLys 41
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Percent Similarity:
                                                                                                                                                                                                                                         WO9928482-A2.
                                                                                                                                                                                                                                                                                                    28-JUL-1998;
28-NOV-1997;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Alignment Scores:
                                                                                                                                                                                                                                                                               24-NOV-1998;
                                                                                                                               07-SEP-1999
                                                                                                                                                                                                                                                                                                                                                     Malcolm BA,
                                                                                                                                                                                                                                                            10-JUN-1999
                                                                                                                                                                                                                       Synthetic
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Pred. No.:
                                                                      RESULT 13
                                                                               AAX803
                                                                                                 合
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ö
                                                 a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              364 GCGCCCCCCGGGGCGCGTTCCTTGACACCATGCACCTGTGGCAGCTCAGACCTTTACTTG 423
                                                                                                                                                                                                                                                                                                                                                                                                                303
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ProLysGlyProValThrGlnMetTyrThrAsnValAspLysAspLeuValGlyTrpGln 101
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           162 AlaGlyHisAlavalGlyIlePheArgAlaAlaValSerThrArgGlyValAlaLySAla 181
                                                                                                                                                          NS4A NS3 complex comprising a central hydrophobic domain of native HCV NS4 Perptide, and an HCV NS3 serine protesse domain, where the hydrophobic domain of native HCV NS4 peptide is terbered by the linker to the amino terminus of the HCV NS3 protease domain. The present sequence encodes an example of the above complex. The covalent NS4A-NS3 complexes are useful for structural determination and determination of mode of binding of HCV inhibitors by NMR spectroscopy. They can also be used for detecting inhibitors of the protease activity the helicase activity and the ATPase activity of NS3. The covalent NS4A-NS3 complexes are more soluble, stable and active than the non-
                                                                                                                                                                                                                                                                                                                                                                                                     5
                                                                                                                                                                                                                                                                                                                                                                                                                                            4.1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          81
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     5 GlySerValValIleValGlyArgIleAsnLeuSerGlyAsp------ThrAlaTyr
                                                                                                                                                                                                                                                                                                                                                                                                                                            22 AlaGlnGlnThrArgGlyGluGluGlyCysGlnGluThrSerGlnThrGlyArgAspLys
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          62 SerileAsnGlyValLeuTrpThrValTyrHisGlyAlaGlyThrArgThrIleAlaSer
                                                                                                                                                   present invention describes a covalent hepatitis C virus (HCV)
                                                                                                                                                                                                                                                             covalent protease-peptide complexes previously available.
                                                                                                                                                                                                                                                                               Sequence 1998 BP; 410 A; 596 C; 568 G; 424 T; 0 other;
                                                                                                                                                                                                                                                                                                             1998
167
15
11
3
                                                                                                                                                                                                                                                                                                             Length:
Matches:
Conservative:
Mismatches:
                                                                                                                                                                                                                                                                                                                                                    Indels:
Gaps:
                                                                      Yao N;
                                                                                                            New hepatitis C virus covalent complexes
                                                                                                                                 Disclosure; Page 160-162; 211pp; English
                                                                                                                                                                                                                                                                                                                                                                                  0S-09-965-594-18 (1-197) x AAX80353 (1-1998)
                                                                      Weber PC,
                                                                                                                                                                                                                                                                                                             7.77e-73
877.50
92.86%
85.20%
98WO-US24528
                    98US-0094331
97US-0067315
                                                                     Taremi SS,
                                                   (SCHE ) SCHERING CORP
                                                                                         WPI; 1999-385385/32
                                                                                                                                                                                                                                                                                                                                            Best Local Similarity:
                                                                                                                                                                                                                                                                                                                                  Percent Similarity:
                                                                                                                                                                                                                                                                                                    Alignment Scores:
                     28-JUL-1998;
28-NOV-1997;
24-NOV-1998;
                                                                      Malcolm BA,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 82
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  424
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          142
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                122
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           484
                                                                                                                                                                                                                                                                                                                                                    Query Match:
OB:
                                                                                                                                                                                                                                                                                                               :.
9
                                                                                                                                                                                                                                                                                                               Pred.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ó
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               à
ò
                                                                                                                                                                                                                                                                                                                                                                                                                       q
                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 õ
```

```
The present invention describes a covalent hepatitis C virus (HCV) NS4A-NS3 complex comprising a central hydrophobic domain of native HCV NS4A-PEPTIGE. and an HCV NS3 serine protease domain, where the hydrophobic domain of native HCV NS3 perine is tethered by the linker to the amino terminus of the HCV NS3 protease domain. The present sequence encodes an example of the above complex. The covalent NS4A-NS3 complexes are useful for structural determination and determination of mode of binding of HCV inhibitors by NMR spectroscopy. They can also be used for detecting inhibitors by the protease activity, the helicase activity and the Arbase activity of NS3. The covalent NS4A-NS3 complexes are more soluble, stable and active than the noncovalent protease-peptide complexes previously available.
HCV; hepatitis C virus; single chain recombinant complex; linker; NS3 protease; NS4a cofactor; NS4a-NS3 complex; serine protease; hydrophobic domain; covalent complex; detection; inhibitor; ss.
                                                                            182 ValAspPheIleProValGluSerLeuGluThrThrMetArgSerPro 197
                                                                                                        Sequence 1998 BP; 410 A; 596 C; 568 G; 424 T; 0 other;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1998
167
15
11
11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Length:
Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                                      HCV NS4A-NS3 complex encoding cDNA SEQ ID NO:104.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Disclosure; Page 163-166; 211pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               New hepatitis C virus covalent complexes
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Yao
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Gaps:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             US-09-965-594-18 (1-197) x AAX80354 (1-1998)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      દ્વે
                                                                                                                                                                                                                                          AAX80354 standard; cDNA; 1998 BP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Weber
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  7.77e-73
877.50
92.86%
85.20%
86.28%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    98WO-DS24528
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       98US-0094331
97US-0067315
                                                                                                                                                                                                                                                                                                                                                   (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (SCHE ) SCHERING CORP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Taremi
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         WPI; 1999-385385/32.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Best Local Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Hepatitis C virus
```

```
483
                                                                                                                                                                                                                                                                        543
                                                                                                                                                                                                                                                                                                       243
                                                                                               303
                                                                                                                                LeuSerProArgProlleSerTyrLeuLysGlySerSerGlyGlyProLeuLgsPro 161
102 AlaProGlnGlySerArgSerLeuThrProCysThrCysGlySerSerAspLeuTyrLeu 121
                                                                                                                                                                                                       ValThrArgHisAlaAspValIleProValArgArgArgGlyAspSerArgGlySerLeu 141
                                                                                                                                                                                                                                                                                             162 AlaGlyHisAlaValGlyIlePheArqAlaAlaValSerThrArqGlyValAlaLysAla 181
                                                                                                                     ProLysGlyProValThrGlnMetTyrThrAsnValAspLysAspLeuValGlyTrpGln 101
                                  19
                                                                            81
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             has
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Hepatitis C; ds; viral prototoxophore; anti-viral; tumour; virus; infection; antitumour; toxophore; human immunodeficiency virus; HIV infection; herpes simplex virus; HSV; rhinovirus; NS3 protease.
                                                                         184 CTCTCCCCCAGGCCTGTCTCCTACTTGAAGGGCTCTTCGGGTGGTCCACTGCTCTGGCCCT
                                AsnGlnValGluGlyGluValGlnIleValSerThrAlaThrGlnThrPheLeuAlaThr
                                                      Novel synthetic viral prototoxophore for treating viral infections, oxim molety incorporated into substrate domain specific for viral enzyme, bound and modified by viral enzyme to get converted into toxophore.
                                                                                                                                                                                                                                                                                                                                       ValAspPheIleProValGluSerLeuGluThrThrMetArgSerPro 197
                                                                                                                                                                                                                                                                                                                                                  Anti-viral synthetic prototoxophore associated DNA sequence.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Shepard
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Example 1; Page 62; 66pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                      BP.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Neuteboom SIC,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  26-APR-2002; 2002WO-US13223.
                                                                                                                                                                                                                                                                                                                                                                                                     ABX15706 standard; DNA; 612
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       27-APR-2001; 2001US-286893P.
                                                                                                                                                                                                                                                                                                                                                                                                                                                28-MAR-2003 (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            (NEWB-) NEWBIOTICS INC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      WPI; 2003-167102/16.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        WO200287500-A2.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Cathers BE,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           07-NOV-2002
                                  42
                                                                                                                     83
                                                                                                                                                                                   364
                                                                                                                                                                                                         122
                                                                                                                                                                                                                                                   142
                                                                                                                                                                                                                                                                                                                                       182
                                                                                                                                                                                                                                                                                                                                                           604
                                                                                                                                                                                                                                                                                                                                                                                                                           ABX15706;
                                                                                                                                                                                                                                                                                                                                                                                 RESULT 14
                                                                                                                                                                                                                                                                                                                                                                                          a
                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                          a
                                                                                                                                                                                g
                                                                                                                                                                                                       ď
                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                               ö
                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                      ò
                                                   8
                                                                    ò
                                                                                         q
                                                                                                                  ŏ
                                                                                                                                   QQ
                                                                                                                                                       ò
                                à
```

This invention relates to a novel synthetic viral prototoxophore

```
comprising a toxin moiety operatively incorporated into a substrate

domain specific for a viral enzyme. This prototoxophore may be bound
and anodified by the viral enzyme. This prototoxophore may be bound
and anodified by the viral enzyme thus converting it to a toxophore.

Condition of the invention is a method for enhancing the anti-viral
effect of an antiviral agent, this method comprises contacting a cell,
infected with a virus or is susceptible to infection, with a
prototoxophore. The invention further comprises an assay to identify
anti-viral agents, comprising contacting an infected cell with a
candidate agent and comparing the ability of the agent to inhibit the
conforting way have virusing the ability of the agent to inhibit the
prototoxophores of the invention may be useful for reducing or
further cell, connective tissue cell, muscle cell (e.g. lymphocyte,
nerve cell, connective tissue cell, muscle cell or hepatocyte) which is
infected with a virus or is susceptible to infection with a virus, with
an effective amount of the prototoxophore. The cells are cell lines
adapted to long term continuous culture or isolated from a subject.
The prototoxophore is also useful for amellorating the severity of a
viral infection in a subject, where the virus is selected from human
immunodeficiency virus (HIV), herpes simplex virus (HSV), rhinovirus and
hepatitis virus, by administering an effective amount of the
prototoxophore to the subject. The prototoxophores of the invention are
also useful for treating tumours. The present sequence represents an
antiviral prototoxophore associated DNA sequence, this sequence is
the invention although in it is clearly not a protein in a protein and a prototoxophore is a recombinant with a protein prototion and antiviral prototoxophore is a recombinant with a protein and a prototion and antiviral prototoxophore is prototoxophore in a prototion and antiviral prototoxophore is prototoxophore in a prototion and antiviral prototoxophore is prototoxophore in a prototoxophore in a pro
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                122 ValThrArgHisAlaAspValileProValArgArgArgGlyAspSerArgGlySerLeu 141
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               438
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            LeuSerProArgProlleSerTyrLeuLysGlySerSerGlyGlyProLeuLeuCysPro 161
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           439 crerceccececcarrrecracrreadagecrecreegegegerecrererecec 498
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           AlaGlyHisAlaValGlyIlePheArgAlaAlaValSerThrArgGlyValAlaLysAla 181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              81
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       GGTAGTGTGTGTGTGGGTAGGATCATTTTGTCCGGTAGTGGTAGTATCACGGCGTAC 78
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      22 AlaGlnGlnThrArgGlyGluGluGlyCysGlnGluThrSerGlnThrGlyArgAspLys 41
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 379 GTCACGAGGCACGCCGATGTCATTCCCGTGCCCGCGGGGTGATAGCAGGGGCAGCCTG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          GlySerValValIleValGlyArgIleAsnLeuSerGlyAsp-----ThrAlaTyr
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            62 SerIleAsnGlyValLeuTrpThrValTyrHisGlyAlaGlyThrArgThrIleAlaSer
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      the invention although it is clearly not a protein sequence.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Sequence 612 BP; 120 A; 171 C; 191 G; 130 T; 0 other;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Conservative:
Mismatches:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Length:
Matches:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (1-612)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             US-09-965-594-18 (1-197) x ABX15706
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      3.39e-73
874.50
91.28%
89.74%
85.99%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Percent Similarity:
Best Local Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       19
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               142
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           162
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Query Match:
DB:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Pred
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      d
         $$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
```

```
completed: August 30, 2003, 19:48:09 e : 188.939 secs
                                                                                                                                                                                          102
                           42
                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Search cou
Job time
                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                           ŏ
                                                 合
                                                                               ò
                                                                                                       g
                                                                                                                                   à
                                                                                                                                                              g
                                                                                                                                                                                          ò
                                                                                                                                                                                                               임
                                                                                                                                                                                                                                              ö
                                                                                                                                                                                                                                                                        셤
                                                                                                                                                                                                                                                                                                õ
                                                                                                                                                                                                                                                                                                                                                       à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       The present invention describes a covalent hepatitis C virus (HCV) NS4A-NS3 complex comprising a central hydrophobic domain of native HCV NS4A-NS3 complex comprising a central hydrophobic domain of native HCV NS4A peptide is tethered by the linker to the amino terminus of the HCV NS4A peptide is tethered by the linker to the amino terminus of the HCV NS3 protease domain. The present Sequence encodes an example of the above complex. The covalent NS4A-NS3 complexes are useful for structural determination and determination of mode of binding of HCV inhibitors by NWR spectroscopy. They can also be used for detecting inhibitors of the protease activity, the helicase activity and the ATPses activity of NS3. The covalent NS4A-NS3 complexes are more soluble, stable and active than the noncovalent protease-peptide complexes previously available.
499 GCGGGGCACGCCGTGGGCATATTTAGGGCCGCGCTGTGCACCCGTGGAGTGGCTAAGGCG 558
                                                                                                                                                                                                                  HCV; hepatitis C virus; single chain recombinant complex; linker; NS3 protease; NS4A cofactor; NS4A-NS3 complex; serine protease; hydrophobic domain; covalent complex; detection; inhibitor; ss.
                           182 ValAspPhelleProValGluSerLeuGluThrThrMetArgSer 196
                                         651
166
16
10
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Sequence 651 BP; 120 A; 187 C; 200 G; 144 T; 0 other,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Conservative:
Mismatches:
Indels:
                                                                                                                                                                                           HCV NS4A-NS3 complex encoding cDNA SEQ ID NO:95
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Length:
Matches:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Yao N;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Disclosure; Page 147-148; 211pp; English
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      New hepatitis C virus covalent complexes
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Weber PC,
                                                                                                            ВР
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            3,66e-73
874.50
93,33%
85,13%
                                                                                                            AAX80345 standard; cDNA; 651
                                                                                                                                                                                                                                                                                                                                                                                              98US-0094331
97US-0067315
                                                                                                                                                                                                                                                                                                                                                                      98WO-US24528
                                                                                                                                                                (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Taremi SS,
                                                                                                                                                                                                                                                                                                                                                                                                                                          (SCHE ) SCHERING CORP.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         WPI; 1999-385385/32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Score:
Percent Similarity:
Best Local Similarity:
                                                                                                                                                                                                                                                                         Hepatitis C virus.
Synthetic.
                                                                                                                                                                                                                                                                                                                  W09928482-A2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                              28-JUL-1998;
28-NOV-1997;
                                                                                                                                                                                                                                                                                                                                                                    24-NOV-1998;
                                                                                                                                                                 07-SEP-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Malcolm BA,
                                                                                                                                                                                                                                                                                                                                            10-JUN-1999
                                                                                                                                      AAX80345;
                                                                                                             셤
```

අ δ

5 GlySerValValileValGlyArglleAsnLeuSerGlyAsp------ThrAlaTyr 21

US-09-965-594-18 (1-197) x AAX80345 (1-651)

Query Match: DB:

AlaGlnGlnThrArgGlyGluGluGlyCysGlnGluThrSerGlnThrGlyArgAspLys 41

```
101
                                                                                                                                                                                                                121
                                                                                                                                                                                                                                                                                                                                                                                                 162 AlaGlyHisAlaValGlyIlePheArgAlaAlaValSerThrArgGlyValAlaLySAla 181
                                                                                                          124 TCCCAACAGACGCGGGGCCTACTTGGTTGCAAGAAGACTAGCTTACAGGCCGGGACAAG 183
                                                62 SerileAsnGlyValLeuTrpThrValTyrHisGlyAlaGlyThrArgThrIleAlaSer
                                                                                                                                                   82 ProLysGlyProValThrGlnMetTyrThrAsnValAspLysAspLeuValGlyTrpGln
                                                                                                                                                                                                                AlabroglnGlySerArgSerLeuThrProCysThrCysGlySerSerAspLeuTyrLeu
                              AsnGlnValGluGlyGluValGloIleValSerThrAlaThrGlnThrPheLeuAlaThr
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          648
                                                                                                                                                                                                                                                                                                                                                                                                                                                           182 ValAspPheIleProValGluSerLeuGluThrThrMetArgSer 196
```

```
GenCore version 5.1.6
Copyright (c) 1993 ~ 2003 Compugen Ltd.
```

```
August 30, 2003, 19:20:43; Search time 1910.31 Seconds (without alignments) 2506.388 Million cell updates/sec
                                                                                                                                                          1 MKKKGSVVIVGRINLSGDTA.....VAKAVDFIPVESLETTMRSP 197
OM protein – nucleic search, using frame_plus_p2n model
                                                                                                                                                                                                                                                                                                                          22781392 seqs, 12152238056 residues
                                                                                                                                                                                                                    Xgapop 10.0 , Xgapext
Ygapop 10.0 , Ygapext
Fgapop 6.0 , Fgapext
Delop 6.0 , Delext
                                                                                                                     US-09-965-594-18
1017
                                                                                                                                                                                                     BLOSUM62
                                                                                                                     Title:
Perfect score:
                                                                                                                                                                                                   Scoring table:
                                                                                                                                                              Sequence:
                                                                                                                                                                                                                                                                                                                          Searched:
                                       Run on:
```

45562784 Total number of hits satisfying chosen parameters: Minimum DB seq length: 0 Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0% Maximum Match 100% Listing first 45 summaries

Command line parameters:
-WODEL-frame+\_p2n.model -DEV=xlp
-WODEL-frame+\_p2n.model -DEV=xlp
-WODEL-frame+\_p2n.model -DEV=xlp
-Worgn2\_1/USFTO\_spool/US0996594/runat\_29082003\_151919\_28322/app\_query.fasta\_1.2872
-DB-EST - OPMT-fastap -SUPFIX=rst -MINMATCH-0.1 -LOOPCL-0 -LOOPEXT=0
-UNITS-b1ts -START=1 -RMP=1.1 -MATRIX=b1csum62 -TRANS-human40.cd1 -LIST=45
-DOCALIGN=200 -THR\_SCORE-pct -THR\_MAX=100 -THR\_MIN-0 -ALIGN=15 -MODE-LOCAL
-UOTFRT-pto -NORM-axt -HEAPSIZE=500 -NHRILN-0 -NAXLEN-2000000000
-USER-US9965594\_eCGN\_1\_1\_15630\_drunat\_29082003\_151919\_28322 -NCPU-6 -ICPU-3
-NO\_MMAP -LARGEQUERY -NEG\_SCORES-0 -WAIT -DSPELOCK=100 -LONGLOG
-DSV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THRRADS-1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6

em\_gss\_pln:\* em\_gss\_vrt:\* em\_gss\_fun:\* em\_gss\_phg:\* em\_gss\_vrl:\* gb\_gssl:\* gb\_est3:\*
gb\_est4:\*
gb\_est5:\*
em\_estfun:\* em\_estom:\* em\_gss\_hum:\* em\_gss\_mus:\* em\_gss\_inv:\* em\_gss\_mam:\* em\_gss\_pro:\* em\_gss\_rod:\* em\_estba:\*
em\_esthum:\*
em\_estin:\*
em\_estin:\*
em\_estpy:\* em\_estro:\* em\_htc:\* gb\_est1:\* gb\_est2:\*
gb\_htc:\* EST:\* Database :

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

		æ			SUMMAKIES	
Result No.	Score		₽.	DB	QI	cription
0 1	; 6		. 7		BMQ15803	04 A
	10			4 ~	000000000000000000000000000000000000000	000
			1404	1 -	B0026101	, כ
	: 5		9	2 -	B0892497	2 5
0			50	90	AA036834	) <del>-</del>
	'n		701	10	BF863244	4
	95.5		984	10	BF304699	BF304699 601888252
	6		629	10	BG089727	27
6 0	93.5		905	13	BU542842	2
Н	ش		1733	12	BM553374	74
11	93		228	12	BM402566	99
-	83		644	29	BX238988	88
~	93		2	29	B2342381	81
c 14	σ		1213	13	BU541777	11
-	ς.		772	29	CC406704	04
16		٠	789	50	CC406705	02
~ c	į,		86.6	13	BQ894657	23
10 O	U. 1.0		87C	2,5	A0620249	4.5
6.4 C.C.	;∘	٠	7 0	7 -	BQ923437	200
212	16		546	: :	BC85398	9 0
12	91		753	3 .	BU402910	10
7	93		998	13	BU219343	43
c 24	6		906	13	BX434207	0
~	0	٠	81	Π.	CNS09179	65
26	90.5	•	1141	Ξ;	AK080545	45
<b>,</b>	ż٩	•	4 0	7 :	BM46/2/9	2, 5
0 00	2 2		3,00	7 -	BM (0500)	36
1 m	86		617	10	BE055917	3 6
31	6		622	i က	AV835401	, ~
c 32	90	•	631	10	AW961059	59
m (	06		658	12	BMB30847	47
M) (	26		757	12	B1258851	25
n ~	200		, os		BAU93694	ž ;
0 30	28		859	3 =	BQ45451 B0222616	17
(r)	6		286	13	BX357170	70 BX357
39	90		0	13	BX404631	31 BX404
C 40	90	•	9	29	CNS040AW	73 Tetra
41	90	•	1123	13	BX398349	49 BX398
42	90	٠	50		AL562877	7 AL5628
. <del>.</del>	06	•	22		BX339469	69 BX339469
4. 4 4. R	26.08	20 a	4004	7.	BGB2/880	BG827880 602752394
•	'n		h •		2	SS AMGNNOC:
					ALIGNMENTS	
100	Ų					
us us	,	3M915803			1146 bp mRNA	linear EST 12-MAR-2002
DEFINITI	~, •	COURT	663945	5 NIH	_MGC_41 Homo sapiens	one IMAGE: 5482
ACCESSIO		15803	orian ha			
VERSION	BM91	5803.	1 GI:19	193661	182	
SOURCE			emid) soc	( ueu		
ORGANI	SM	o sapiens	Sus	Ì		
	Euk Kan	Eukaryota; Mammalia:	; Metazoa; Entheria:	, ag	Chordata, Craniata, Vertebrata Primates, Catarrhini, Dominida	brata; Euteleostomi;
REFERENCI	ω	(bases	1 to 11		Cacatiniin	ı.

a

```
147 eSerTyrLeuLysGlySerSerGlyGlyProLeuLeuCysProAlaGlyHisAlaValGl 167
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CB950999
CB950999.1 GI:30205777
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 3.8
104.50
41.10%
31.29%
10.28%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1. 1031
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Percent Similarity:
Best Local Similarity:
                                                                                                                                                                                                                                186 oval 187
                                                                                                                                                                                                                                                                                       681 TTTG 678
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        235
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Query Match:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            source
                                                                                                                                                                                                                                                                                                                                                                                                                          DEFINITION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Pred. No.:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     BASE COUNT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ORGANISM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         VERSION
KEYWORDS
SOURCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ACCESSION
                                                                                                                                                                                                                                                                                                                                                                      CB950999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              FEATURES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ORIGIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                a
                                                          g
                                                                                                                    ö
                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                       a
                                                                                                                                                                                                                                å
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                // Abortism = nature of the first series / mol_type="marks" | // white = nature | // w
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            949
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     859
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           90 rThrAsnValAspLysAspLeuValGlyTrpGlnAlaProGln------GlySerAr 107
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          107 9SerLeuThrProCysThrCysGlySerSerAspLeuTyrLeuValThrArgHisAlaAs 127
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         127 pValileProValArgArgArgGlyAspSerArgGlySerLeuLeuSerProArgProIl 147
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 61 rSerIleAsnGlyValLeuTrpThrValTyrHis------GlyAlaGlyThr-- 76
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                833 -------CGGGAGGCGCGCGGGGCCCGGTAGTGTGGGCCCGGGTCCCCTCT 787
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         22 AlaGlnGlnThrArgGlyGluGluGlyCysGlnGluThrSer-GlnThrGlyArgAspLy 41
                                                                                                        Email: cgapDS-remail.nih.gov
Tissue Procurement: DCTD/DTP
CDNA Library Preparation: Rubin Laboratory
CDNA Library Preparation: Rubin Laboratory
CDNA Library Paraged by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MCC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
http://image.llnl.gov
lld.gov
High quality sequence start: 6
High quality sequence stop: 256.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 902 ------CAGATGTGCGGGTGGGAAGCGCCCGCTCGCGCGGGGGGCCAG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            77 -------ArgThrIleAlaSerPro------LysGlyProValThrGlnMetTy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    948 AGGGGAGGGGCGCGGGGGGGGCGCGGGGGGGGGGGCCCG------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        41 sAsnGlnValGluGlyGluValGlnIleValSerThrAlaThrGlnThrPheLeuAlaTh
                                                                                                                                                                                                                                                          ğ
                                National Institutes of Health, Mammalian Gene Collection (MGC)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1146
47
19
59
57
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              858 ACTIGCTIGICGIGITITCIGIGGG-------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Length:
Matches:
Conservative:
Mismatches:
Indels:

    1146
    /organism="Homo sapiens"

                                                          Unpublished
Contact: Robert Strausberg, Ph.D.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Gaps:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       US-09-965-594-18 (1-197) x BM915803 (1-1146)
     NIH-MGC http://mgc.nci.nih.gov/
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          3.1
106.00
36.26%
25.82%
10.42%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Percent Similarity:
Best Local Similarity:
Query Match:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 169
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Alignment Scores:
AUTHORS
TITLE
JOURNAL
COMMENT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              BASE COUNT
ORIGIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Pred. No.:
                                                                                                                                                                                                                                                                                                                                                                                                                                FEATURES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             đ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ç
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ద
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        õ
```

```
CB950999

AGENCOORT_13445496 NIH_MGC_177 Mus musculus cDNA clone
AMMAGE:30316162 5', mRNA sequence.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           167 yilepheArgAlaAlaValSerThrArgGlyValAlaLySAlaValASpPhe---IlePr 186
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 1031)
HH-MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  44 ValGluGlyGluValGlnIleValSerThrAlaThrGlnThrPheLeuAlaThrSerIle 63
                                                                         þe
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CDNA Library Preparation: Michael Brownstein Laboratory CDNA Library Arrayed by: The I M.A.G.E. Consortium (LLNL) DNA Sequencing by: Apencourt Bioscience Corporation Clone distribution: MGC clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Plate: NDCMIO7 row: b column: 11
High quality sequence stop: 333.
Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        /organism="Mus musculus"
/mol_type="mRNA"
/db_xref="texon:10090"
/clone="IMAGE:30316162"
/lab_host="DH108 (T1-phage-resistant)"
/clone_lib="NIH_MGC_177"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1031
51
16
62
35
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Contact: Robert Strausberg, Ph.D.
Email: cgapbs-rémail.nih.gov
Tissue Procurement: Dr. Michael Brownstein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Length:
Matches:
Conservative:
Mismatches:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          US-09-965-594-18 (1-197) x CB950999 (1-1031)
                                                                                                                                                                                                                                                                                                                                                                           Mus musculus (house mouse)
```

BASE COUNT 297 a 521 c 237 g 345 t 3 others ORIGIN	Alignment Scores: Pred. No.: Score: Score: Score: Pred. No.: 102.50 Matches: Percent Similarity: 35.50% Conservative: 12 Best Local Similarity: 29.50% Mismatches: 67 Ouery Match: 10.08% Indels: 62 DB:	-09-965-594-18 (1-197) x BQ926101 (1-1403)  11 GlyArglleAsnLeuSerGlyAspThrAlaTyrAlaGlnGlnThrArgGlyGluGlyGluGly	31 Cysclaglufhr	Db 1249 GGGGTATGGTTATCACGGGCTGGGCAGGTACTTCCCCTAAAGCG 1205  Qy 84 lyprovalThrGlnMetTyrThrAsnvalaspLysAspLeuvalGlyTrpGlnAlaProG 104		Qy         155 GlyGlyProLeuCys 160           Db         979 GCATTCCGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG	
64 AsnGlyValLeuTrpThrValTyrHisGlyAlaGlyThrArgThrIleAlaSerProLys 83	Qy 84 GlyProvalThrGlnMetTyrThrasnValaspLysAspLeuValGlyTrgLnAlara 493  Qy 84 GlyProvalThrGlnMetTyrThrasnValaspLysAspLeuValGlyTrgLnAlaro 103	119 LeuTyrLeuValThrArgHisAlaAsp-ValIleProValArgArgArgGlyAspSerAr 13   1   1   1   1   1   1   1   1   1	bb 652 AAACNCCTTAACGCTCCCCACAACAATTTCTTCAAGGGTCCCTGATTCCTTAA 711  Oy 156 -GlyProLeuLeuCysProAlaGlyHisAlaValGlyIlePheArgAlaAlaValSerTh 175  Db 712 GTCCCCCTTTGTTACCCAGACCACTTTGTGGGACACAAGGCCTTCTTTATTTTTTTT	QY         195 gSerPro 197           b         799 AACACCC 805           RESULT 3         80926101           LOCUS         BQ926101           1403 bp         πRNA           LOCUS         BQ926101           1403 bp         mRNA           LOCUS         BQ926101           LOCUS         BQ926101 <td>5', mRNA sequence. B0926101 B0926101.1 G1:22341132 EST. Mus musculus (house mouse) M Mus musculus Eukaryota; Metazoa; Chordata; Crai</td> <td>REFERENCE 1 (Dases 1 to 1403) AUTHORS NIH-MGC http://mgc.nci.nih.gov/. AUTHORS NIH-MGC http://mgc.nci.nih.gov/. TILLE National Institutes of Health, Mammalian Gene Collection (MGC) JOURNAL Unpublished Cohect: Robert Strausberg, Ph.D. Email: cgapbs-r@mail.nih.gov Tissue Procurement: Mark Maconochie, Ph.D. and Nancy L. Freeman, Ph.D.</td> <td>CDNA Library Preparation: ResGen, Invitrogen Corp CDNA Library Arrayed by: THAIN.4 CE. Consortium (LLML) DNA Sequencing by: Agencourt Bioscience Corporation Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov Plate: LLAM13798 row: j column: 07 High quality sequence stop: 101. FEATURES Location/Qualifiers Location/Qualifiers    I. 1403     Amb.bot.="mana" mna"     Amb.bot.="mana" mna"     Amb.bot.="mana"     Amb.bot.="mana"     Alb.bot.="minde: 6335718"     Alb.bot.="minde: 6335718"     Alb.bot.="minde: 61000"     Antel: EcoRV: Site.2: NotI: Cloned unidirectionally.   Site.1: EcoRV: Site.2: NotI: Cloned unidirectionally.   Primer: Oligo dT: Average insert size 1.95 kb.   Constructed by ResGen, Invitrogen Corp. Note: this is a</td>	5', mRNA sequence. B0926101 B0926101.1 G1:22341132 EST. Mus musculus (house mouse) M Mus musculus Eukaryota; Metazoa; Chordata; Crai	REFERENCE 1 (Dases 1 to 1403) AUTHORS NIH-MGC http://mgc.nci.nih.gov/. AUTHORS NIH-MGC http://mgc.nci.nih.gov/. TILLE National Institutes of Health, Mammalian Gene Collection (MGC) JOURNAL Unpublished Cohect: Robert Strausberg, Ph.D. Email: cgapbs-r@mail.nih.gov Tissue Procurement: Mark Maconochie, Ph.D. and Nancy L. Freeman, Ph.D.	CDNA Library Preparation: ResGen, Invitrogen Corp CDNA Library Arrayed by: THAIN.4 CE. Consortium (LLML) DNA Sequencing by: Agencourt Bioscience Corporation Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov Plate: LLAM13798 row: j column: 07 High quality sequence stop: 101. FEATURES Location/Qualifiers Location/Qualifiers    I. 1403     Amb.bot.="mana" mna"     Amb.bot.="mana" mna"     Amb.bot.="mana"     Amb.bot.="mana"     Alb.bot.="minde: 6335718"     Alb.bot.="minde: 6335718"     Alb.bot.="minde: 61000"     Antel: EcoRV: Site.2: NotI: Cloned unidirectionally.   Site.1: EcoRV: Site.2: NotI: Cloned unidirectionally.   Primer: Oligo dT: Average insert size 1.95 kb.   Constructed by ResGen, Invitrogen Corp. Note: this is a

```
AA036834 50.2 bp mRNA linear EST 26-AUG-1996 zk29d05.rl Soares_pregnant_uterus_NbHPU Homo sapiens cDNA clone AMCE:471945 5' similar to PTR:A55195 A55195 chordin precursor -African clawed frog ;, mRNA sequence.
                                                                                                                                                                                                                        Email: estewatson.wustl.edu
This clone is available royalty-free through LLNL; contact the
This clone is available royalty-free through LLNL; contact the
TMAGE Consortium (info@image.llnl.gov) for further information.
Seq primer: -2BM13 rev2 from Amersham
High quality sequence stop: 300.
Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      34 ThrSerGlnThrGly-----ArgAspLysAsnGlnValGluGlyGluValGlnIleVal
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        287 CTGACTCTTCTGGGAACCACTGCCCAGCAAAACGGCAGCCCCGGGGCCCATC---
                                                                                                                                                                                                                                                                                                                                                                                             Washington University School of Medicine 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108 Tel: 314 286 1800
Fax: 314 286 1810
       871 TCAGGGCGTTTTAAAGCCCCCGGCCTTCGCGCCGGGCGAAGCA 913
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1. .502
/organism="Homo sapiens"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             /mol_type="mRNA"
/db_xref="GDB:3757947"
/db_xref="taxon:9606"
/clone="IMAGE:471945"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  US-09-965-594-18 (1-197) x AA036834 (1-502)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    /sex-"female"
                                                                                                                                                          AA036834.1 GI:1509872
                                                                                                                                                                                            Homo sapiens (human)
Homo sapiens
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             11.6
95.50
38.51%
25.00%
9.39%
                                                                                                                                                                                                                                                                                                                                                               Unpublished
Contact: Wilson RK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Percent Similarity:
Best Local Similarity:
Query Match:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             84
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Alignment Scores:
                                                       AA036834/c
LOCUS
DEFINITION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      source
                                                                                                                                                                          KEYWORDS
SOURCE
ORGANISM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Pred. No.:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             BASE COUNT
                                                                                                                                                                                                                                                                                                                                             TITLE
JOURNAL
COMMENT
                                                                                                                                          ACCESSION
VERSION
                                                                                                                                                                                                                                                            REFERENCE
AUTHORS
                                          RESULT 5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    FEATURES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ORIGIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ö
                                                                                                                                                                                                                                                                                                                          /clone_lib="Lupaki_sympathetic_trunk"
/clone_lib="Lupaki_sympathetic_trunk"
/note="Vector: pcMv-SPORT6 (Life Technologies); Site_l:
/Not1: Site_2: Sall; cDNA made by oligo-dT priming.
Directionally cloned using the following adaptors:
- TrGACCACGCGTCGA 3' and
5'-TGACACACACGCGCGCGCCCT(15)-3'. Size selected >
1 kb for average insert length 1.9 kb. This is a primary
Ikb for average insert length 1.9 kb. This is a primary
Itbrary, non-amplified. Library constructed by Life
Technologies and donated by J. Lupski, M.D./Ph.D. (Baylor
College of Medicine); available through Life
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             132
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           143 ------SerProArgProIlcSerTyrLeuLys-------GlySerSerGl 155
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                751 AAAACCCACCTCGGCCCACCGATGCGCTAAGCCTCCCTTTACAAGCCACCGCGCGGG 810
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  yGlyProLeuLeuCysProAlaGlyHisAlaValGlyIlePheArgAlaAlaValSerTh 175
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  484 TGGGATCCATTTTAATAAAGGGTGCTCTTTAATCATGGCGCCCACGGGCCCGGTGATA 543
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        88 GlnMetTyrThrAsnValAspLysAspLeuValGlyTrpGlnAlaProGlnGlySerArg 107
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      68 TrpThrValTyrHisGlyAlaGlyThrArgThrIleAlaSerProLysGlyProValThr 87
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Blossience Corporation
Clone distribution: MGC Clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM13595 row: column: 13
High quality sequence start: 57
High quality sequence start: 57
High quality sequence start: 67
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           119 LeuTyr-LeuValThr-----ArgHisAlaAspValIleProValArg----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        28 others
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      rArgGlyValAlaLysAlaValAspPheIleProValGluSer 189
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1199
40
17
54
6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Length:
Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                        /tissue_type="sympathetic trunk"
/dev_stage="adult, 16 yr"
/lab_host="DH10B"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 211 t
                                                                                                                                                     Location/Qualiflers
1 1199
/organism="Homo sapiens"
/mol_type="mRNA"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Gaps:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    US-09-965-594-18 (1-197) x BQ892487 (1-1199)
                                                                                                                                                                                                                   /db_xref-"taxon:9606"
/clone-"IMAGE:6192708"
/sex-"male"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   343 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Technologies."
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               11.5
100.50
36.77%
25.81%
9.88%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Percent Similarity:
Best Local Similarity:
Query Match:
DB:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   255
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  155
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      175
                                                                                                                                                                            source
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               BASE COUNT
ORIGIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ..
₽
                                                                                                                                                          FEATURES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Score:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Pred
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  음
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
```

Not

502 37 20 45 46

Length:

--- 237 21

S

```
Percent Similarity:
                     Alignment Scores:
Pred. No.:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        150
                                                                                                             Query Match:
DB:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            RESULT 7
BF304699/c
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                DEFINITION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ORGANISM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           REFERENCE
AUTHORS
TITLE
JOURNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ACCESSION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      VERSION
KEYWORDS
SOURCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           FEATURES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        COMMENT
                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                a
                                                                                                                                                                                                                                                                                                                                                    õ
                                                                                                                                                                                                                                                                                                                                                                                    q
                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                             đ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        엄
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1...(Organism-"chlamydomonas reinhardtii"
/mol_type="mRNA"
/stran="cc.1890 wild type mt+ 21gr"
/db_xref="taxon:3055"
/dboxel="taxon:3055"
/clone_lib="C. reinhardtii CC-1690, Stress condition I,
normalized, Lambda Zap II"
/note="Vector: pBluescript II SK-: Site_l: ECORI; Site_2:
XhOI; This library, constructed by John Davies and Jeffrey
McDemnott, combines cDNAs from CC-1690 cells grown to
mid-log phase in TAP-N (30 min, lhr, 4hr), TAP-S (30 min,
lhr, 4hr), TAP-P (4hr, 12hr, 24hr), NO3 to NH4 (30min, lhr,
4hr) and NH4 to NO3 (30min, lhr, 4hr), PalyA mRNA was
purified from each sample, pooled and cDNA synthesized.
The cDNA was directionally cloned into lambda Zap II
(Stratagene) in the Ecorn (5) and XhOR (3') sites.
pBluescript II SK- plasmids were excised from the lambda
ZAP clones by superinfection with ExAssist (Stratagene)
phage- The library was normalized using method 4 described
in Bonaldo et al (1996) Genome Research 6: 791-806."
                                                                                                                                                                                                                                                                                                                                                                                                                                                                              BF863244 10-JAN-2001 963042C02.xl C. reinhardtii CC-1690, Stress condition I, normalized , Lambda 2ap II Chlamydomonas reinhardtii CDNA, mRNA sequence.
                                                                                                                                                                                                                                                                                            83 LysGlyProValThrGlnMetTyrThrAsnValAspLysAspLeuValGlyTrpGlnAla 102
                                                                                                                                                                      103 ProGlnGlySerArgSerLeuThrProCysThrCysGlySerSerAspLeuTyrLeuVal 122
                     38
                                                                        82
                                                                                                                                                                                                                                                                                                                               77 ACAGCACACTTA-----ATTAAGCCAAAGGGGGCACAACGGGGTGCCNANCN 30
SerThrAlaThrGlnThrPheLeuAlaThrSerIleAsnGlyValLeu------ 67
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Chlamydomonas reinhardtii
Chlamydomonas reinhardtii
Eukaryota: Viridiplantae: Chlorophyta; Chlorophyceae; Volvocales;
Chlamydomonadaceae; Chlamydomonas.
1 (bases 1 to 701)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Grossman, A., Davies, J., Federspiel, N., Harris, E., Bauser, C., Lefebvre, P., McDermott, J. P., Shraqer, J., Sliflow, C. and Stern, D. Analyses of the Chlamydomonas reinhardtii Genome: A Model, Unicellular System for Analysing Gene Function and Regulation in Vascular Plants; project phase 3
                                                                                                                                                                                                                                                      137 AGCCGGGGACACTGCACCTTCTCACAGTGCACCTCTCCAGTGCCCCCTTGCAGGTGCAG
                                                                        .------TrpThrValTyrHisGlyAlaGlyThrArgThrIleAlaSerPro
                                                                                             chauser@duke.edu.
Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                     142 uSerProArgProIleSerTyr 149
                                                                                                                                                                                                                                                                                                                                                                                            GTACCCGTTGCCCGCCAGCTTC 8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  BF863244
BF863244.1 GI:12253388
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Contact: Charles Hauser
DCMB Box 91000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Duke University
Durham, NC 27708-1000
Tel: 919 613 8159
Fax: 919 613 8177
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Unpublished
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          173
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Email:
                                      236
                                                                        9
                                                                                                                                                                                 167
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SOURCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Source
                                                                                                                                                                                                                                                                                                                                                                                                                                                                              LOCUS
DEFINITION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        BASE COUNT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       JOURNAL
COMMENT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ACCESSION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     VERSION
KEYWORDS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              REFERENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               AUTHORS
                                                                                                                                                                                                                                                                                                                                                                                                                                          RESULT 6
BF863244
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     FEATURES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    TITLE
                                                                                                                                                                            g
                                                                                                                                                                                                                      à
                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                           8
                                                                                                                                                                                                                                                                                                                                                                                              셤
                                    a
                                                                        å
                                                                                                   g
                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                               ö
```

ORIGIN

```
BF304699 984 bp mRNA linear EST 21-NOV-2000 6018888252F1 NIH_MGC_17 Homo sapiens cDNA clone IMAGE:412276 5',
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     464
                                                                                                                                                                                                                                                                      171 CACCACCATACCCTTGCTCTCACGTTGCTCACCAAAATTATGCCCATACGGGCCACTA 230
                                                                                                                                                                                                                                                                                                                                                                                                                                                   404
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              170 ArgAlaAlaVal-------181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               525 AAAGAAACCATTTGAGTAGGAAGTCGCCCGTTTACCCCCGAAGGTGAAGGTCACCTCTAT 584
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. I (bases 1 to 984)
NIH-WGC http://mgc.noi.nih.gov/.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
Plate: LLCM1005 row: g column: 13
High quality sequence stop: 646.
Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                  107 ArgSerLeuThrProCysThrCysGlySerSerAspLeuTyrLeuValThrArgHisAla
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  139 -----Tyr
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   87 ThrGlnMetTyrThrAsnValAspLysAspLeuValGlyTrpGlnAlaProGlnGlySer
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 345 GAGATCATTIGCAIGIGGCTTIAGTCACCCCAAGAGAGCCCIGGGAGIGGGCATITAIAA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                LeuLysGlySerSerGlyGlyProLeuLeuCysProAlaGlyHisAlaValGlyIlePhe
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              465 CTCCAAGGCAGCAATGGGAGCCTTTCGCGGTGTGCGGGTGCTCGTCGTCTTAATGTCAGG
                                                                                                                                                                                                                                                                                                                                                                    231 ACAAAGTTACATACACGG -----AAGGACCAGCGCGCTTGGCCACCCCCTTGGAGCG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      182 ValAspPheIleProValGluSerLeuGluThrThrMetArg 195
                                                                                                                                                                                                                      71 TyrHisGlyAlaGlyThrArgThrIleAlaSerProLys----
     701
38
22
22
63
31
  Length:
Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  /organism="Homo sapiens"/mol_type="mRNA"
                                                                                                                            Gaps:
                                                                                                                                                                        US-09-965-594-18 (1-197) x BF863244 (1-701)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         mRNA sequence.
BF304699
BF304699.1 GI:11251586
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Homo sapiens (human)
Homo sapiens
17.9
95.50
38.96%
24.68%
9.39%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Unpublished
                                                                         Best Local Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           source
```

Sequencing Center information can be

9

```
1. 629
/organism="Mus musculus"
/mol_type="mark"
/mol_type="mark"
/db_xref="taxon:10090"
/clone="IMAGE:3977578"
/tlsaue_type="Mx calls (flow-sorted)"
/lab_host="bHl0B (Il-resistant)"
/clone_lib="NCI_CGAP_SP2"
/note="Organ: splen: Vector: pcMv-SPORT6 (Life
/note="organ: splen: Vector: pcMv-SPORT6 (Life
Technologies); mRNA made from flow-sorted Nx cells, cDNA
made by oligo-dT priming. Directionally cloned. Average
insert size 1.5 kb. primary library, non-amplified. cDNA
Library Preparation: David B. Krizman, Ph.D."
156 c 150 g 191 t l others
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                91 ThrAsnValAspLysAspLeuValGlyTrpGlnAlaProGlnGlySerArgSerLeu--- 109
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ----GlyAspSerArg 138
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               608 GAAGGAGACACACAGATCATCCCTTGTGCA---AAAACATTTCCCATTCCAGATATAAAT 552
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 159 LeucysProAlaGlyHisAlaValGlyIlePheArgAlaAla-------ValSer 174
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         500 AAGAGAACTAAAGCTGTGAGACCCCTCAAGTTGCCCAGATCCAATGCCCGGGTGAAGCCA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               :::|||
233 GTGTGTGAC---AACCAAGCATATGGACTTTTCGCCTATGCAAAAAACGGAACAATCTCT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            45 GluGlyGluValGlnIleValSerThrAlaThrGlnThrPheLeuAlaThrSerIleAsn
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 74 ------AlaGlyThrArgThrIleAlaSerProLysGlyProValThrGlnMetTyr
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         440 GNGAATGTG---TGCAGTGTGGCTGGCTGG------GGGTCAAGGTCCATCAAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  110 ---ThrProCysThrCysGlySerSerAspLeuTyrLeuValThrArgHisAlaAspVal
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          395 GACACTAAAGCATCTGCCCCCCTGCGAGAGGTTCAACTGGTCATCCAGGAGGACGAGGAA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         335 TGCAAAAAACGTTTCCGATACTACACTGAGACCACAGAGATTTGTGCTGGAGACTTGAAG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           275 ---AAAATAAAGACTCCT-------TTCAAGGGTGACTCGGGGGGACCCCTT
  DNA Sequencing by: Washington University Genome Clone distribution: NCI-CGAP clone distribution found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    629
46
22
28
58
46
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   175 ThrArgGlyValAlaLysAlaValAspPhellePro 186
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        65 GlyValLeuTrpThrValTyrHisGly-------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Length:
Matches:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Gaps:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   129 IleProvalArgArgArg------
                                                                                                                                                                       Seq primer: -400P from Gibco
High quality sequence stop: 422.
Location/Qualiflers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 US-09-965-594-18 (1-197) x BG089727 (1-629)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 17.5
95.00
39.53%
26.74%
9.34%
                                                                                                                                               MGI:1477610
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Best Local Similarity:
Query Match:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        131
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Percent Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     BASE COUNT
ORIGIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    RESULT 9
BU542842/c
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Pred. No.:
                                                                                                                                                                                                                                   FEATURES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        d
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      à
                 /clone=TMAGE:4122256"
/tissue_type="rhabdomyosarcoma"
/tissue_type="rhabdomyosarcoma"
/tissue_type="rhabdomyosarcoma"
/tissue_type="rhabdomyosarcoma"
/tissue_type="rhabdomyosarcoma"
/tissue_type="rhabdomyosarcoma"
/tissue_type="rhabdomyosarcoma"
/clone_lib="rhallom"
/clone_lib="rhallom"
/clone_lib="rhallom"
/clone_tion=lip"
/clone_tip="rhabdomyosarcom"
/clone_tip="rhabdomyosarcom"
/clone_tip="rhabdomyosarcom"
/clone_tip="rhabdomyosarcom"
/clone_tip="rhabdomyosarcomyosarcomyosarcomyosarcomyosarcomyosarcomyosarcomyosarcomyosarcomyosarcomyosarcomyosarcomyosarcomyosarcomyosarcomyosarcomyosarcomyosarcomyosarcomyosarcomyosarcomyosarcomyosarcomyosarcomyosarcomyosarcomyosarcomyosarcomyosarcomyosarcomyosarcomyosarcomyosarcomyosarcomyosarcomyosarcomyosarcomyosarcomyosarcomyosarcomyosarcomyosarcomyosarcomyosarcomyosarcomyosarcomyosarcomyosarcomyosarcomyosarcomyosarcomyosarcomyosarcomyosarcomyosarcomyosarcomyosarcomyosarcomyosarcomyosarcomyosarcomyosarcomyosarcomyosarcomyosarcomyosarcomyosarcomyosarcomyosarcomyosarcomyosarcomyosarcomyosarcomyosarcomyosarcomyosarcomyosarcomyosarcomyosarcomyosarcomyosarcomyosarcomyosarcomyosarcomyosarcomyosarcomyosarcomyosarcomyosarcomyosarcomyosarcomyosarcomyosarcomyosarcomyosarcomyosarcomyosarcomyosarcomyosarcomyosarcomyosarcomyosarcomyosarcomyosarcomyosarcomyosarcomyosarcomyosarcomyosarcomyosarcomyosarcomyosarcomyosarcomyosarcomyosarcomyosarcomyosarcomyosarcomyosarcomyosarcomyosarcomyosarcomyosarcomyosarcomyosarcomyosarcomyosarcomyosarcomyosarcomyosarcomyosarcomyosarcomyosarcomyosarcomyosarcomyosarcomyosarcomyosarcomyosarcomyosarcomyosarcomyosarcomyosarcomyosarcomyosarcomyosarcomyosarcomyosarcomyosarcomyosarcomyosarcomyosarcomyosarcomyosarcomyosarcomyosarcomyosarcomyosarcomyosarcomyosarcomyosarcomyosarcomyosarcomyosarcomyosarcomyosarcomyosarcomyosarcomyosarcomyosarcomyosarcomyosarcomyosarcomyosarcomyosarcomyosarcomyosarcomyosarcomyosarcomyosarcomyosarcomyosarcomyosarcomyosarcomyosarcomyosarcomyosarcomyosarcomyosarcomyosarcomyosarcomyosarcomyosarcomyosarcomyosarcomyosarcomyosarcomyosarcomy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      BG089727 629 bp mRNA linear EST 26-JAN-2001 mab90e06.x1 NCI_CGAP_Sp2 Mus musculus cDNA clone IMAGE:3977578 3' similar to SW:GRAD_MOUSE P11033 GRANZYME D PRECURSOR;, mRNA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           100 TrpGlnAlaProGlnGlySerArgSerLeuThr---ProCysThrCysGlySerSerAsp 118
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                119 LeuTyrLeuValThrArgHisAlaAspValIleProValArgArgArgGlyAspSerArg 138
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           139 GlySerLeuLeuSerProArgProIleSerTyrLeuLysGlySerSerGlyGlyProLeu 158
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       159 LeuCysProAlaGlyHis-Ala-----ValGlyIlePheArg-----AlaAlaVa 173
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              553 GGCGCCTCTTG-------527
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      173 lSerThrArgGlyValAlaLySAlaValAspPheIleProValGluSerLeuGluThrTh 193
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalla; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 629)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Contact: Robert Strausberg, Ph.D.
Email: cgapbs-rémail.nih.gov
Tissue Procurement: David Segal Ph.D., Herbert Morse M.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               984
984
34
74
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Length:
Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Gaps:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           US-09-965-594-18 (1-197) x BF304699 (1-984)
/db_xref="taxon:9606"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Mus musculus (house mouse)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      BG089727.1 GI:12572290
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      427 CAAGGGGGCGCCC 415
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     193 rMetArgSerPro 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      27.9
95.50
42.86%
37.14%
9.39%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Tumor Gene Index
Unpublished
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Mus musculus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Percent Similarity:
Best Local Similarity:
Query Match:
DB:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   sednence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           BG089727
                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   869
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      LOCUS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 RESULT 8
BG089727/c
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Pred. No.:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      AUTHORS
TITLE
                                                                                                                                                                                                                                                                                                                                                                             COUNT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              REFERENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             JOURNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           COMMENT
                                                                                                                                                                                                                                                                                                                                                                          BASE CO
```

8

рp

ŏ a ò

å

ద à 셤 õ

ò

a

158

441

501

```
/do.me="INAGE:742981"
/tissue_type="medula"
/tissue_type="medula"
/lab host="DH108"
/cloom_lib="With McC_119"
/force_lib="With McC_119"
/note="Organ: brain; Vector: pCMV-SPORT6; Site_1: Not1;
/note="Organ: brain; Vector: pCMV-SPORT6; Site_1: Not1;
/note="Organ: brain; Vector site is oligo-dT primed and anonymous male age 27. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.3 kb, insert size range 0.9-3 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 013. Note: this is a NIH MGC Library.

43 a 673 c 521 g 288 t 8 others
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1733 bp mRNA linear EST 20-FEB-2002 AGENCOURT_6558368 NIH_MGC_119 Homo sapiens cDNA clone IMAGE:5742981 BM553374
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Contact: Robert Strausberg, Ph.D.

Email: cgapba-rémail.nih.gov

Tissue Procurement: Life Technologies, Inc.

CDNA Library Preparation: Life Technologies, Inc.

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Agencourt Biscience Corporation

Clone distribution: MGC Clone distribution information can be found through the I.M.A.G.E. Consortium/LLML at:

http://image.lnl.gov

Plate: LLAM12761 row: p column: 22

High quality sequence start: 88

High quality sequence start: 88

High quality sequence stop: 539.

Location/Qualifiers

I. 1733

rce

//db.xref**Homo Sapiens**
//db.xref**Taxon:5210061**
                                                                                                                                                                                                                                                                                                                                       584 GCACTGCCGCCCAGCCCTCCATCTCAGCGGGATGTGCAGGGTGAGACAGGAATGCAGGGA 525
                                                                                                                                                                              100 TrpGlnAlaProGlnGlySerArgSerLeuThrProCysThrCysGlySerSerAspLeu 119
                                                                                                                                                                                                                                      524 CGTTCTGCCCCTAGGTCAGCCTCTTCATCCGCCTGTTGTGCTTGTGGATGGTCAAGGTTG 465
                                                                                                                                                                                                                                                                                                 120 -----TyrLeuValThrArgHisAlaAsp-----ValIleProValArgArg 133
                                                                                                                                                                                                                                                                                                                                                                                                                     134 Arg------Arg 145
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        146 ProlleSerTyrLeuLysGlySerSerGlyGlyProLeuLeuCysProAlaGly----- 163
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       644 CAGTAGCAGCCCCATCCAGAGGAAGACCACTCCGGAGGGCCACAGGCCTCTGCAGGCCTTG 585
                                                           80 AlaSerProLysGlyProValThrGlnMetTyrThrAsnValAspLysAspLeuValGly 99
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Eukaryota, Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalla; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1733)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  NIH-MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        164 -----HisAlaValGly 167
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   284 CGGGTGGCTCACGGTGCAGGG 264
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          BMS53374.1 GI:18792049
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Homo sapiens (human)
Homo sapiens
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Unpublished
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           243 8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            RESULT 10
BM553374/c
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           BASE COUNT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   DEFINITION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ORGANISM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     AUTHORS
TITLE
JOURNAL
COMMENT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ACCESSION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         REFERENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          VERSION
KEYWORDS
SOURCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      FEATURES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                g
                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                         염
                                                                                                                                                                                                                                                                                                                                                                                                                        à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                           à
                                                                                                                                                                                                                                                                                                                                                                                                     Contact: Robert Strausberg, Ph.D.

Email: cgapbs-rémail.nih.gov
Tissue Procurement: DCTD/OTP

CDNA Library Preparation: Ruin Laboratory

CDNA Library Preparation: Ruin Laboratory

CDNA Library Preparation: Ruin Laboratory

CDNA Library Preparation: Ruin Laboration

CDNA Library Preparation: Ruin Laboration

CLONG distribution: Information can be

found through the I.M.A.G.E. Consortium/LINL at:

Location/Qualifiers

Location/
BU542842 905 bp mRNA linear EST 13-SEP-2002 AGENCOURT_10334715 NIH_MGC_40 Homo sapiens cDNA clone IMAGE:6574789
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           04 AAGIGGGAGGCAGGCGATGGTGCACCTGGGGAGGCCCCTGGTGCAGAAGCAGCCCCA 645
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          28 GluGluGlyCysGlnGluThrSerGln - - - ThrGlyArgAspLysAsnGlnValGluGly 46
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            70 -------ValTyrHisGlyAla------74
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         -----GlyThrArgThrile 79
                                                                                                                                                                                                                                   Eukaryota, Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalla; Eutheria; Primates; Catarrhini; Hominidae; Homo. 1 (bases 1 to 905) NIH-MGC http://mgc.nci.nih.gov/. NIH-MGC http://mgc.nci.nih.gov/. National Institutes of Health, Mammalian Gene Collection (MGC)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     47 GluValGlnIleValSerThrAlaThrGlnThrPheLeuAla----------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Length:
Matches:
Conservative:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Mismatches:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          US-09-965-594-18 (1-197) x BU542842 (1-905)
                                                           5', mRNA sequence.
BU542842
BU542842.1 GI:22853325
                                                                                                                                                                                 Homo sapiens (human)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            39.5
93.50
31.88%
24.64%
9.19%
                                                                                                                                                                                                             Homo sapiens
                                                                                                                                                                                                                                                                                                                                                                                         Unpublished
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Best Local Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         75 -----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Percent Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Query Match:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     source
LOCUS
DEFINITION
                                                                                 ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Pred. No.:
                                                                                                                                                                                                                                                                                              REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            BASE COUNT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   FEATURES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ORIGIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          9
```

JOURNAL Unpublished Contact: Cushman JC Contact: Cushman JC Department of Biochemistry University of Nevada MS200, Reno, NV 89557-0014, USA Tel: 775-784-1050 Email: jcushmaneunr.edu	FORMARD: 72 Juner BACKWARD: 72 Juner BACKWARD: 72 Juner Plate: 005 row: F column: 12 Seq primer: 73 Juner Location/Qualifiers 1. 528 Location/Qualifiers 2. 500	according to manufacture's (Stratusgene, Inc.) recommended protocol for the Lambda UnizapxR vector and cDNA synthesis kit."  BASE COUNT 129 a 125 c 137 g 137 t and cDNA synthesis alignment Scores:  Pred. No.: 21.9 Length: 528 Score: 93.00 Matches: 37 Percent Similarity: 42.98 Conservative: 15 Best Local Similarity: 30.58 Mismatches: 43 Query Match: 9.14 Indels: 26 DB: 12 Gaps: 4 US-09-965-594-18 (1-197) x BM402566 (1-528)	Oy 94 AspLysAspLeuvalGlyTrpGlnalaProGlnGlySerArgSerLeuThrProCysThr 113	Oy         139 GlyserLeuLeuSerProArgPro11eSerTyrLeu         150           Db         230TCAGCCGCTAATGCTCGTCCAATCCAGACGACGACGCTATTAAT         286           Qy         151 LysGlySerSerGlyGlyProLeuLeuCysProAlaGlyHisalavalGlyIlePhearg         170           I(1:::
92.2 92.2 37.26 arity: 26.424 1.19	US-09-965-594-18 (1-197) x BM553374 (1-1733)  QY	Qy         80 AlaSerProLysGlyProValThrGlnMetTyrThrAshValAspLeuValGly 99           1         ::::::::::::::::::::::::::::::::::::	Oy 152 GlySerSerGlyGlyProLeuLeuCysProAlaGlyHisAlaValGlyIlePheArgAla 171    11	RESULT 11 BM402566 BM402560 SIAONOFFL2.34513 An expressed sequence tag (EST) collection from the resurrection plant Selaginella lepidophylla Selaginella RECESSION BM402566.1 GI:21643782 KEYWORDS SOURCE Solaginella lepidophylla ORGANISM BUKATYOTA: Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Lycopodlophyta; Isoetopsida; Selaginellales; Selaginellaceae; Selaginella Chases I to 528) AUTHORS TITLE An expressed sequence tag (EST) collection from the resurrection

σ

```
/organism="Sorghum bicolor"
/wol_type="genomic DNA"
/wol_type="genomic DNA"
/wol_type="genomic DNA"
/done='ic83bil"
/lab_host="JM107 or DH5a"
/clone='ic83bil"
/clone='ib='wGS-SbicolorF (JM107 adapted methyl filtered)"
/clone='ib='wGS-SbicolorF (JM107 adapted methyl filtered)"
/clone='ib='wGS-SbicolorF (JM107 adapted methyl filtered)"
/note="Site_l: Xba I; Site_2: Xba I; The vector was
digested with Xba In one nucleotide was added by fill in
in the recessive 3' end. The genomic DNA was nebulized,
end repaired, adaptor ligated and size fractionated using
sephadex. The resulting fragments were between 0.8 and 3
kb and were cloned into the vector (xx/y reads in M13mpl9,
-b/g reads in pUC19: The same ligation was transformed in
either JM107 or DH5a."
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     113
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ------CysGlySerSerAsp 118
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    119 LeuTyrLeuvalThrArgHisAlaAspValIleProValArgArgArgGlyAsp----- 136
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             320 CACCCAGTGCCCCCGGCGGCCGCCGTACGGGGACAGGAAACTGGTGTGTGAGGATCAC 261
                                                                                                                         Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermarothyta; Wagnollophyta: Lillopsida; Poales; Poaceae; PacCAD clade; Panicoideae; Andropogoneae; Sorghum.

1 (bases 1 to 701)
Rablnowicz,P. D. O'Shaughnessy,A. L., Balija,V., Dedhia,N., Katzenburger,F., King,L., Miller,B., Muller,S., Nascimento,L., Zutavern,T., Palmer,L., McCombie,W.R. and Martienssen,R.A. Genomic shotgun sequences from Sorghum bicolor (methyl-filtered)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              440 GIGGCCGIGGCCCGCCCCCCCCCCCCCCGCGAATTAIGGCCGCAATAAACTCGGCAC 381
ic83b11.bl WGS-SbicolorF (JM107 adapted methyl filtered) Sorghum
bicolor genomic clone ic83b11 5', genomic survey sequence.
BZ342381
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 60 AlaThrSerIleAsnGlyValLeuTrpThrValTyrHisGlyAlaGlyThrArgThr---
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         79 -----IleAlaSerProLysGlyProValThrGlnMetTyrThrAsnValAspLysAsp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     LeuValGlyTrpGlnAlaProGlnGlySerArg---SerLeuThrProCysThr-----
                                                                                                                                                                                                                                                                                                             Contact: W. Richard McCombie
Contact: W. Richard McCombie
Lita Annenberg Hazen Genome Sequencing Center
Cold Spring Harbor Laboratory
PO Box 100, Cold Spring Harbor, NY 11724, USA
Tel: 516 367 8884
Email: mccombieGeshl.org
Plate: ic83 row: b column: 11
Seq primer: -21M13UnivPwd
Class: shotgun
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 701
48
14
56
51
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Length:
Matches:
Conservative:
Mismatches:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                High quality sequence stop: 701
Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       US-09-965-594-18 (1-197) x BZ342381 (1-701)
                                                                                                 Sorghum bicolor (sorghum)
                                                         BZ342381.1 GI:24742983
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 31.7
93.00
36.69%
28.40%
9.14%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1. .701
                                                                                                                       Sorghum bicolor
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Best Local Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        108
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Percent Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       97
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      114
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Query Match:
DB:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           source
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Pred. No.:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        BASE COUNT
  DEFINITION
                                                                                                 SOURCE
                                                                                                                                                                                                                                                                            TITLE
JOURNAL
COMMENT
                                                                                                                                                                                                  REFERENCE
AUTHORS
                                         ACCESSION
                                                         VERSION
KEYWORDS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      FEATURES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ORIGIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ద
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ô
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    GSS 06-NOV-2002
                                                                                                                                                                                                                                                                        Submitted (27-JAN-2003) The Sanger Institute, Wellcome Trust Genome Campus, Hinxton, Cambridgeshire, CBIO 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk Unpublished
This sequence was generated from the T7 end of BAC 283L13. 283L13 is part of the Daniokey BAC Library created by R. Plasterk and N.V. Keygene. Further details:
                    GSS 29-JAN-2003
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        345 AGCAGCTGGTCATGTCGGTGGTCCAGCAGTCTTGGAAAGGAC------TGGAGATCC 295
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              294 AGACCGTIGGGGACAAAGGCGGAGGAGTCGTCACCACCATGCCCTTTAAAAAAGGAGAAAA 235
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     LeuTyrLeuValThrArgHisAlaAspValileProValArgArgArgArgGlyAspSerArg 138
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          234 TIATITGIGATIACCATGGGGAGGAGITGTCGCACGAGGAGGCCTTAGGGAGACGGGAGG 175
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    139 GlySerLeuLeuSerProArgPro-----IleSerTyrLeuLysGlySerSerGlyGly 156
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                157 ProLeuLeu---CysProAlaGlyHisAlaValGlyIlePheArgAlaAlaValSerThr 175
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ---ProGlnGlySerArg-----SerLeuThrProCysThrCysGlySerSerAsp 118
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           GlnMetTyr------ThrAsnValAspLysAspLeuValGlyTrpGlnAla 102
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       28
                                                                                                                                                        117 CCCTTTGCCTGGATGCCCAGTCCTGGCCCTGCCCCTGTCACCCGGGCATGGAAACCTTCG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      TrpThrValTyrHisGlyAlaGlyThrArgThrIleAlaSerProLysGlyProValThr
        ыхдзвувы 644 bp DNA linear GSS 29-JAN-
Danio rerio genomic clone DKEY-283Ll3, genomic survey sequence.
ВX238988
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    linear
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          644
37
118
52
116
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                              http://www.sanger.ac.uk/Projects/D_rerio/
Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                        /mol_type="genomic DNA"
/db_xref="taxon:7955"
/clone="DKEY-283L13"
/tissue_type="Testis"
/note="vector pindigoBAC-536"
a 212 c 176 g 127 t
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    DNA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Length:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Gaps:

    .644
    /organism-"Danio rerio"

                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    701 bp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              US-09-965-594-18 (1-197) x BX238988 (1-644)
                                                                              BX238988.1 GI:28161322
                                                                                                                     Danio rerio (zebrafish)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        28.4
93.00
44.72%
30.08%
9.14%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               176 ArgGlyVal 178
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      GAAGGCTTA 49
                                                                                                                                        Danio rerio
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Percent Similarity:
Best Local Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    BZ342381
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         119
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      68
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               103
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Query Match:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            BASE COUNT
ORIGIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                BZ342381/c
                  LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              .. ON
                                                                                                                                                                                                                                     AUTHORS
TITLE
JOURNAL
  8X238988/c
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             RESULT 13
                                                                                                                                                                                                                         REFERENCE
                                                                                                                                                                                                                                                                                                                                                                                                                   FEATURES
                                                                                                                                                                                                                                                                                                                                          COMMENT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    COCUS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     å
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             õ
```

ô

```
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
  US-09-965-594-18 (1-197) x BU541777 (1-1213)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      /organism-"Zea mays"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Location/Qualifiers
1. .772
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Email: whitelaw@tigr.org
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Other_GSSs: PUHKD12TD
Contact: Cathy Whitelaw
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CC406704.1 GI:30886794
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Class: sheared ends
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Seq primer: TR
                                                                                                                                                                                                                                                                                                                                                       114 -----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            180 LysAla 181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            11 CGCTCT 6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    2ea mays
2ea mays
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   150
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   source
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   BASE COUNT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CC406704/c
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    DEFINITION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ORGANISM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   TITLE
JOURNAL
COMMENT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ACCESSION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 REFERENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      AUTHORS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 VERSION
KEYWORDS
SOURCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           FEATURES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                  a
                                                                                                                                                     ò
                                                                                                                                                                                                  d
                                                                                                                                                                                                                                                       à
                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                       õ
                                                                                                                                                                                                                                                                                                                                                                                                       셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                            ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 /MD_xref="txxon:9606"
/dD_xref="txxon:9606"
/clone="IMAGE:6573516"
/tisxoue_type="carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_40"
/clone_lib="NIH_MGC_40"
/clone_lib="NIH_MGC_40"
/clone_lib="NIH_MGC_40"
/clone_lib="tx"
/c
                                                                                                                                                                                                                                                                                                                                                                                                                                                                          AGENCOURT_10325396 NIH_MGC_40 Homo sapiens cDNA clone IMAGE:6573516 5', mENA sequence.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Contact: Robert Strausberg, Ph.D.

Email: cgapbs-r@mail.nih.gov
Tissue Procurement: DCTD/DTP

CDNA Library Preparation: Rubbin Laboratory

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Agencourt Bloscience Corporation

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:

http://mage.lln.gov

Plate: LLCM2767 row: f column: 12

High quality sequence stop: 165.
                                                                             150 LeuLysGly-----LeuCy 160
                                                                                                                                                                                                                                                       160 sproAlaGlyHisAlaValGlyIlePheArgAlaAlaValSerThrArgGlyVal---- 178
GIGGAGAGGGCAGGACGACGACGA-------CGACGACGACGGGGGGACCCCC 216
                                                     ---SerArgGlySerLeuLeuSerProArgProIleSerTyr 149
                                                                                                                                                                                                  155 TITACIGCICGGCCTCGGCCCCGGCCTCGGCCGCGGGGGGGGCCGTTGGATGIG 96
                                                                                                                                                                                                                                                                                  Eukaryofa; Metazoa: Chordata: Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria: Primates: Catarrhini; Hominidae; Homo. (bases 1 to 1213) NIH-MGC http://mgc.nci.nih.gov/. National Institutes of Health, Mammalian Gene Collection (MGC)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1213
41
13
50
50
8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Length:
Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1. 1213
/organism-"Homo sapiens"
/mol_type-"mRNA"
                                                                                                                                                                                                                                                                                                                                                       179 -AlaLysAlaValAspPheIlePro 186
                                                                                                                                                                                                                                                                                                                                                                                                       47 rececrecerrecrerecre 23
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         BU541777.1 GI:22852258
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Homo sapiens (human)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                64.9
93.00
38.03%
28.87%
9.14%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Homo sapiens
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Unpublished
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Percent Similarity:
Best Local Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Alignment Scores:
260
                                                  137
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Query Match:
DB:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            VERSION
KEYWORDS
SOURCE
ORGANISM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         source
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            LOCUS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Pred. No.:
                                                                                                                                                                                                                                                                                                                                                                                                                                                          RESULT 14
BU541777/c
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                BASE COUNT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ACCESSION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              FEATURES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ORIGIN
                                                                                                                                                                                                                                                                                                                                                                                                    ద
                                                                                                  셤
                                                                                                                                                                                          ద
                                                                                                                                                                                                                                                                                                                                                    ò
                                                     ð
                                                                                                                                               ö
                                                                                                                                                                                                                                                       ç
                                                                                                                                                                                                                                                                                                    셤
```

```
CC406704 772 bp DNA linear GSS 19-MAY-2003 WHENIZTB ZM_0.6_1.0_KB Zea mays genomic clone ZMMBTa469B24, genomic survey sequence.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Illiopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea . 1 (bases 1 to 772)
Whitelaw, C. A., Quackenbush, J., Van Aken, S., Utterback, T., Resnick A., Fraser, C.M., Yuan, Y., San Miguel, P., Ma, J. and Bennetzen, J. Maize Genomics Consortium
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   /db_xref="taxon:e>/,
/clone="zMmBra469B24"
/clone_11b="zM_0.6_1.0_KB"
/note="Vector: pCR4_TOP0; Site_1: EcoR1; 0.6-1.0 kb high
CoT selected genomic DNA library"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          162 AlaGlyHisAlaVal-----GlyIlePheArgAlaAlaValSerThrArgGlyValAla 179
                                                                                                                                                                                                                                                                                             -----CysGlySerSerAspLeuTyrLeuValThrArgHisAla 126
                                                                                                                                                                                                                                                                                                                                                                                                                                                 127 AspValileProValArgArgArg-------GlyAspSerArgGlySerLeu 141
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      142 LeuSerProArgProlleSerTyrLeuLysGlySerSerGlyGlyProLeuLeuCysPro 161
                                                            401 IGGGCCGIGCAGICCACITIAGCCACAIAGACITIGGCAICIICC ------AIG 354
                                                                                                                        GloMetTyrThrAsnValAspLysAspLeuValGly------TrpGlnAlaPro--- 103
                                                                                                                                                                                     353 TIGITGIATITGICTCGCAGGICATICCAAGTCGGCIGCAICCGCIGGCAGTGICCACAC 294
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       122 CGCCGCCCCATGGCCCAGCAGCAGCAGCAGCGCGCAGTCAGGGCCG------CCG 72
68 TrpThrValTyrHisGlyAlaGlyThrArgThrIleAlaSerProLysGlyProValThr
                                                                                                                                                                                                                                                     104 GinGlySerArgSerLeuThrProCysThr-------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                9712 Medical Center Drive, Rockville, MD 20850, USA Tel: 301-838-5843
Fax: 301-838-0208
```

ORIGIN			
Aliqument Scores:			
Pred. No.:	40.3	Length:	772
Score:	92.50	Matches:	47
Percent Similarity:	40.578	Conservative:	24
Best Local Similarity:	26.86%	Mismatches:	29
Ouery Match:	9.10%	Indels:	46

29	10	
DB: 29 US-09-965-594-18 (1-197) x CC406704	Gaps:	(1-772)
DB: US-09-965-594-18	29	(1-197) x CC406704
	DB:	US-09-965-594-18

ArgaspLysasnGlnvalGluGlyGluValGlnIleValScrThralaThrGlnThr 57     :::       :::	58 PheLeualaThrSerIleAsnGlyValLeuTrpThrValTyrHisGlyAlaGlyThr 76 :::	ThrileAlaSerProLysGlyProValThrGlnWetTyrThrAsnValAspLysAsp 96	537 CAGACTITGCTGACGAIGAGATCAAGGGTGACACAGAIGAIGACCAGGAAGAGC 484	97 LeuValGlyJrpGlnAlaProGlnGlySerArgSerLeuThrProCysThr 113	ACCCAGTGGTTTCCCTACATGCAATAACTGGGATTAGAAGGGAGGACACCATGCAGC 424	CysGlySerSerAspLeuTyrLeuValThrArgHisAlaAsp 127	TGCGGGTRGTCCTCAACGCTCAGGAGCTGCTGGCCCTACTTGACACGGGTTCAACACATA 364	128 ValileProValArgArgGlyAspSerArgGlySerLeuLeuSerProArgProIle 147	TCATCAACTGCAAGGCGGCA-CAACAGCGTGGGGTTACTTTGGAACCCACCAGGT 305	148 SerTyrLeuLysGlySerSerGlyGlyProLeuLeuCysProAlaGlyHisAlaValGly 167	.catgicaagtggcaaatggagacccagtititigc	IlePheArgAlaAlaValSerThrArgGlyValAlaLysAlaValAspPhe 184	GTAACTCGTCGCGCA	185	TTCACCATTGAGGCATATGCAATTCCCTTGGATACATTTGAG 179
39 ArgA      633 AGAT	58 PheL	77 ArgT	37 CAGA	97 LeuV	83 AGGA	114 CysG	423 TGCG	28 Vali	63 ACTT	48 SerT	.04 CGCC	168 IleP	259 GTAA	.85	23 AAGT
Oy Db 6	oy da		S da	ογ	Db da	0y 1	Db 4	Qy 1	Dp 3	Qy 1	Dp 3	oy 1	Db 2	0y 1	Dp g

Search completed: August 31, 2003, 04:27:41 Job time : 1917.31 secs

```
AAB15224;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              RESULT 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                AAB15224
                                                                                                                                                                                                         August 30, 2003, 17:42:58 ; Search time 44.6227 Seconds (without alignments) 700.745 Million cell updates/sec
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1. (SIDSI)/gogdata/geneseq/geneseqp-embl/AA1980.DAT:*
2. (SIDSI)/gogdata/geneseq/geneseqp-embl/AA1981.DAT:*
3. (SIDSI)/gogdata/geneseqy-embl/AA1981.DAT:*
4: (SIDSI)/gogdata/geneseqy-embl/AA1981.DAT:*
5. (SIDSI)/gogdata/geneseqy-embl/AA1982.DAT:*
6: (SIDSI)/gogdata/geneseqy-embl/AA1982.DAT:*
7: (SIDSI)/gogdata/geneseqy-embl/AA1986.DAT:*
8: (SIDSI)/gogdata/geneseqy-embl/AA1980.DAT:*
8: (SIDSI)/gogdata/geneseqy-geneseqp-embl/AA1980.DAT:*
8: (SIDSI)/gogdata/geneseqy/geneseqp-embl/AA1980.DAT:*
8: (SIDSI)/gogdata/geneseqy/geneseqp-embl/AA1990.DAT:*
11: (SIDSI)/gogdata/geneseqy/geneseqp-embl/AA1991.DAT:*
12: (SIDSI)/gogdata/geneseqy/geneseqp-embl/AA1991.DAT:*
13: (SIDSI)/gogdata/geneseqy/geneseqp-embl/AA1991.DAT:*
14: (SIDSI)/gogdata/geneseqy/geneseqp-embl/AA1991.DAT:*
15: (SIDSI)/gogdata/geneseqy/geneseqp-embl/AA1991.DAT:*
16: (SIDSI)/gogdata/geneseqy/geneseqp-embl/AA1991.DAT:*
17: (SIDSI)/gogdata/geneseqy/geneseqp-embl/AA1991.DAT:*
18: (SIDSI)/gogdata/geneseqy/geneseqp-embl/AA1999.DAT:*
18: (SIDSI)/gogdata/geneseqy/geneseqp-embl/AA1999.DAT:*
20: (SIDSI)/gogdata/geneseqy/geneseqp-embl/AA1999.DAT:*
21: (SIDSI)/gogdata/geneseqy/geneseqp-embl/AA1999.DAT:*
22: (SIDSI)/gogdata/geneseqy/geneseqp-embl/AA1999.DAT:*
23: (SIDSI)/gogdata/geneseqy/geneseqp-embl/AA2000.DAT:*
24: (SIDSI)/gogdata/geneseqy/geneseqp-embl/AA2000.DAT:*
24: (SIDSI)/gogdata/geneseqy/geneseqp-embl/AA2000.DAT:*
24: (SIDSI)/gogdata/geneseqy/geneseqp-embl/AA2000.DAT:*
                                                                                                                                                                                                                                                                                                                                                                                                                                      1 MKKKGSVVIVGRINLSGDTA......VAKAVDFIPVESLETIMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1107863
GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Total number of hits satisfying chosen parameters:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1107863 seqs, 158726573 residues
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries
                                                                                                                                                OM protein • protein search, using sw model
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          BLOSUM62
Gapop 10.0 , Gapext 0.5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                A_Geneseq_19Jun03:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         DB seq length: 0
DB seq length: 200000000
                                                                                                                                                                                                                                                                                                                                                      US-09-965-594-20
1020
                                                                                                                                                                                                                                                                                                                                                                     Title:
Perfect score:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Scoring table:
                                                                                                                                                                                                                                                                                                                                                                                                                                         Sequence:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Searched:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Database
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Minimum I
Maximum I
                                                                                                                                                                                                                    Run on:
```

		rus	rus	rus	rus	rus	rus	rus	rus	
	tion	tis C vi	Hepatitis C virus	tis C vi						
	Description	Hepati	Hepati	Hepati	Hepati	Hepati	Hepati	Hepati	Hepati	
IES								_		
SUMMARIES	ΩI	AAB15224	AAB15225	AAB15223	AAB15222	AAB15221	AAB15226	AAB15220	AAB15212	
	DB	21	21	21	21	21	21	21	21	ć
	Length	197	197	197	197	197	197	195	195	222
ď	Query Match	100.0	0.66	98.2	97.1	95.4	92.7	92.1	89.4	0 0 0
	Query Score Match Length DB II	1020	1010	1005	066	973	946	686	912	200
	Result No.		71	m	4	Ŋ	9	7	œ	c

HCV NS4A-NS3 COMPL HCV NS4A-NS3 COMPL	NS4A-NS3	HCV NS4A-NS3 COMPI HCV NS4A-NS3 COMPI	NS4A-NS3	HCV NS4A-NS3 COMP1		NS4A-NS3	NS4A-NS3	NS4A-NS3	NS4A-NS3	HCV NS4A-NS3 compl	HCV NS4A-NS3 compl	HCV NS4A-NS3 COMP1	HCV NS4A-NS3 COMP1	HCV NS4A-NS3 compl			NS4A-NS3		HCV NS4A-NS3 compl	43	Hepatitis C virus	Hepatitis C virus	Amino acid sequenc	Hepatitis C virus	HCV genomic amino	pHCV150-encoded se	pHCV176-encoded se	Sequence encoded 1	Protein sequence o	Peptide encoded by	Sequence encoded 1		Peptide encoded by
20 AAY24947 20 AAY24942		20 AAY24941	AAY1788			Ť		20 AAY17883		·	·		·			•		20 AAY17885		20 AAW93482				23 AAU99289			AAR7922		AAP9015		Ī	AAP9205	10 AAP90288
665																					_			~	_	_	m			_	301	2436	436
86.5 86.4		86.0				٠.	٠.	•					85.0		85.0	٠.	٠.					84.2			٠.	•	e.	Э.	۳,	ë.	۳.	•	m.
882.5 881.5	78.	77.	75.	4.	4	73.	73.	7	9.	70.		^		867.5		86	<u>.</u>		S	82	58.	858.5	58.	58.	55.	54.	54.	54.	54.	54.	54.	54.	4
11	12	14	15	16	17	18	6	20	21	22	23	24	25	56	27	28	58	30	31	32	33	34	35	36	37	38	33	04	41	42	43	44	45

## AL IGNMENTS

```
Hepatitis; NS3 protease; viral replication; chronic liver disease; liver failure; liver cancer; mutant; mutein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Goldfarb V;
                                                                                                                                                                                                                                                                                                                          Hepatitis C virus NS4A-NS3 fusion protease #6.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Wittekind M, Weinheimer S, Zhang Y,
AAB15224 standard; protein; 197 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            (BRIM ) BRISTOL-MYERS SQUIBB CO.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   06-JAN-2000; 2000WO-US00345.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      990S-0115271.
                                                                                                                                                                                                             19-DEC-2000 (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                WPI; 2000-465976/40.
N-PSDB; AAA73333.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Hepatitis C virus
Synthetic.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          WO200040707-A1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                08-JAN-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             13-JUL-2000.
                                                                                                            PAR X PAR X
```

Modified hepatitis C virus (HCV) NS3 protease comprising at least 1

```
ö
                                                                                                                                                                                          The present sequence is a mutated version of a fusion protein created using the Hepatitis C virus (HCV) NG3 and NS4A protease enzymes. These proteins are both essential for the replication of the virus, acting to cleave its replicative proteins from the polyprotein produced from the mitviral treatments of the two proteins should be effective as antiviral treatments of HCV infection. This is useful as HCV can lead to chronic liver disease such as cirrhosis, liver failure and liver cancer. The present invention concerns a number of NS3 mutants and NS3-NS4A fusion proteins which can be used to identify inhibitors of this type, as well as enabling structural guidles of the protease and protease: Inhibitor complexes. This sequence contains the alpha-helixo-7
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  61 TSINGVLWTVYHGAGTRTIASPKGPVTQMYTNVDKDLVGWQAPQGSRSLTPCTCGSSDLY 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              121 LVTRHADVIPVRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVSTRGVAK 180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    MKKKGSVVIVGRINLSGDTAYAQQTRGEQGCQKTSHTGRDKNQVEGEVQIVSTATQTFLA 60
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                substitution of a hydrophobic alpha-helix 0 amino acid to a hydrophilic amino acid, useful for screening inhibitors that may treat hepatitis C
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Repatitis; NS3 protease; viral replication; chronic liver disease;
liver failure; liver cancer; mutant; mutein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    .;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              100.0%; Score 1020; DB 21; Length 197; 100.0%; Pred. No. 3.2e-98; tive 0; Mismatches 0; Indels 0;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Goldfarb V;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Hepatitis C virus NS4A-NS3 fusion protease #7.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Zhang Y,
                                                                                                                                  Claim 23; Fig 16; 66pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           AAB15225 standard; protein; 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (BRIM ) BRISTOL-MYERS SQUIBB CO.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                AVDFIPVESLETIMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Wittekind M, Weinheimer S,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              06-JAN-2000; 2000WO-US00345.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             99US-0115271
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          19-DEC-2000 (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Best Local Similarity 100.0
Matches 197; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               WPI; 2000-465976/40.
N-PSDB; AAA73334.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  197 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        WO200040707-A1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             08-JAN-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           13-JUL-2000.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Synthetic.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Sedneuce
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              AAB15225;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         variant
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        AAB15225
AAB
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            õ
```

```
ò
                                                                                                          The present sequence is a mutated version of a fusion protein created using the Hepatitis C virus (HCV) NS3 and NS4A protease enzymes. These proteins are both essential for the replication of the virus, acting to cleave its replicative proteins from the polyprotein produced from the HCV genome. Inhibitors of the two proteins should be effective as antiviral treatments of HCV infection. This is useful as HCV can lead to chronic liver disease such as cirrhosis, liver failure and liver cancer. The present invention concerns a number of NS3 mutants and NS3-NS4A fusion proteins which can be used to identify inhibitors of this type, as well as enabling structural studies of the protease and proteins.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     61 ISINGVLMIVYHGAGIRIASPKGPVIQMINVDKDLVGWQAPQGSRSLIPCTCGSSDLY 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   TSINGVLWTVYHGAGTRTIASPKGPVTQMYTNVDKDLVGWQAPQGSRSLTPCTCGSSDLY 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  LVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVSTRGVAK 180
                                                                                                                                                                                                                                                                                                                                                                                                                                        1 MKKKGSVVIVGRINLSGDTAYAQQTRGEQGCQKTSHTGRDKNQVEGEVQIVSTATQTFLA 60
Modifled hepatitis C virus (HCV) NS3 protease comprising at least 1 substitution of a hydrophobic alpha-helix 0 amino acid to a hydrophilic amino acid, useful for screening Inhibitors that may treat hepatitis C
                                                                                                                                                                                                                                                                                                                                                                                             Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Hepatitis; NS3 protease; viral replication; chronic liver disease; liver failure; liver cancer; mutant; mutein.
                                                                                                                                                                                                                                                                                                                                                                                         ő
                                                                                                                                                                                                                                                                                                                                                         Juery Match
99.0%; Score 1010; DB 21; Length 197;
9est Local Similarity 99.5%; Pred. No. 3.6e-97;
Atches 196; Conservative 0; Mismatches 1; Indels 0;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Weinhelmer S, Zhang Y, Goldfarb V;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Hepatitis C virus NS4A-NS3 fusion protease #5.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Ä
                                                                                  Claim 23; Fig 17; 66pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              AAB15223 standard; protein; 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       (BRIM ) BRISTOL-MYERS SQUIBB CO.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                AVDFIPVESLETIMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   AVDFIPVESLETIMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      990S-0115271
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       06-JAN-2000; 2000WO-US00345
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      WPI; 2000-465976/40.
                                                                                                                                                                                                                                                                                                                           197 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       N-PSDB; AAA73332
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      WO200040707-A1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Wittekind M,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      08-JAN-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           19-DEC-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       13-JUL-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Synthetic
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             AAB15223;
                                                                                                                                                                                                                                                                                                                             Sequence
                                                                                                                                                                                                                                                                                                                                                          Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               181
                                                                                                                                                                                                                                                                                              variant
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               RESULT 3
   õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
```

```
·
0
                                                                                                                                                                                                                                                  The present sequence is a mutated version of a fusion protein created using the Hepatitis C virus (HCV) NS3 and NS4A protease enzymes. These proteins are both essential for the replication of the virus, acting to cleave its replicative proteins from the polyprotein produced from the HCV genome. Inhibitors of the two proteins should be effective as antiviral treatments of they unfection. This is useful as HCV can lead to chronic liver disease such as cirrhosis, liver failure and liver cancer. The present invention concerns a number of NS3 mutants and NS3-NS4A fusion proteins which can be used to identify inhibitors of this type, as well as enabling structural studies of the protease and protease:inhibitor complexes. This sequence contains the alpha-helix0-1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              121 LVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVSTRGVAK 180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          9
                          Modified hepatitis C virus (HCV) NS3 protease comprising at least 1 substitution of a hydrophobic alpha-helix 0 amino acid to a hydrophilic amino acid, useful for screening inhibitors that may treat hepatitis C
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1 MKKKGSVVIVGRINLSGDTAYAQQTRGEQGCQKTSHIGRDKNQVEGEVQIVSTATQTFLA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            TSINGVLWTVYHGAGTRTIASPKGPVTOMYTNVDKDLVGWOAPQGSRSLTPCTCGSSDLY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Hepatitis; NS3 protease; viral replication; chronic liver disease;
liver failure; liver cancer; mutant; mutein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Length 197;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Welnheimer S, Zhang Y, Goldfarb V;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                98.5%; Score 1005; DB 21; 98.5%; Pred. No. 1.2e-96;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Hepatitis C virus NS4A-NS3 fusion protease #4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  2; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Z
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                (BRIM ) BRISTOL-MYERS SQUIBB CO.
                                                                                                                                                                                             Claim 23; Fig 15; 66pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                AAB15222 standard; protein; 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             AVDFIPVESLETTMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           AVDFIPVESLETTMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 99US-0115271
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  06-JAN-2000; 2000WO-US00345
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           19-DEC-2000 (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Local Similarity
nes 194; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 197 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   WO200040707-A1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Wittekind M,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 08-JAN-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  13-JUL-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Synthetic
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   AAB15222;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Best Loca
Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     AAB15222

XX

AAC

AAB

AAC

AAC

AAB

AAC

AAC

AAB

AAC

AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           RESULT 4
q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
```

WPI; 2000-465976/40.

```
ö
                                                                                                                                       The present sequence is a mutated version of a fusion protein created using the Hepatitis C virus (HCV) NS3 and NS4A protease enzymes. These proteins are both essential for the replication of the virus, acting to cleave its replicative proteins from the polyprotein produced from the HCV genome. Inhibitors of the two proteins should be effective as antiviral treatments of HCV infection. This is useful as HCV can lead to chronic liver disease such as cirrhosis, liver failure and liver cancer. The present invention concerns a number of NS3 mutants and NS3-NS4A fusion proteins which can be used to identify inhibitors of this type, as well as enabling structural studies of the protease and
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  LVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVSTRGVAK 180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 121 LVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVCTRGVAK 180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   61 TSINGVLWTVYHGAGTRTIASPKGPVTQMYTNVDKDLVGWQAPQGSRSLTPCTCGSSDLY 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    9
                           Modified hepatitis C virus (HCV) NS3 protease comprising at least 1 substitution of a hydrophobic alpha-helix 0 amino acid to a hydrophilic amino acid, useful for screening inhibitors that may treat hepatitis C
                                                                                                                                                                                                                                                                                                     protease:inhibitor complexes. This sequence contains the alpha-helix0-1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1 MKKKGSVVIVGRINLSGDTAYAQQTRGEQGCQKTSHTGRDKNQVEGEVQIVSTATQTFLA
                                                                                                                                                                                                                                                                                                                                                                                                                         Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Hepatitis; NS3 protease; viral replication; chronic liver disease;
liver failure; liver cancer; mutant; mutein.
                                                                                                                                                                                                                                                                                                                                                                                                                       .:
0
                                                                                                                                                                                                                                                                                                                                                                                         Length 197;
                                                                                                                                                                                                                                                                                                                                                                                                                       4; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Goldfarb V;
                                                                                                                                                                                                                                                                                                                                                                                       Score 990; DB 21;
Pred. No. 4.4e-95;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Hepatitis C virus NS4A-NS3 fusion protease #3.
                                                                                                                                                                                                                                                                                                                                                                                                                    2; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Zhang Y,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Ź
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (BRIM ) BRISTOL-MYERS SQUIBB CO.
                                                                                                            Claim 23; Fig 14; 66pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             AAB15221 standard; protein; 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  181 AVDFIPVESLETTMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                     97.18;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Weinheimer S,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            06-JAN-2000; 2000WO-US00345
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             99US-0115271
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            19-DEC-2000 (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                     Matches 191; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                      Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                       197 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Hepatitis C virus
N-PSDB; AAA73331
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            WO200040707-A1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Wittekind M,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             08-JAN-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Synthetic.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               AAB15221;
                                                                                                                                                                                                                                                                                                                                                        Sequence
                                                                                                                                                                                                                                                                                                                                                                                         Query Match
                                                                                                                                                                                                                                                                                                                          variant
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 AAB15221
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 å
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  å
```

```
Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      RESULT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                    ö
                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   셤
                                                                                                                                                                                                                                                                                                ö
                                                                                                        The present sequence is a mutated version of a fusion protein created using the Hepatitis C virus (HCV) NS3 and NS4A protease enzymes. These proteins are both essential for the replication of the virus, acting to cleave its replicative proteins from the polyprotein produced from the HCV genome. Inhibitors of the two proteins should be effective as antiviral treatments of HCV infection. This is useful as HCV can lead to chronic liver disease such as cirrhosis, liver failure and liver cancer. The present invention concerns a number of NS3 mutants and NS3-NS4A fusion proteins which can be used to identify inhibitors of this type, as well as enabling structural studies of the protease and protease:inhibitor complexes. This sequence contains the alpha-helix0-1
                                                                                                                                                                                                                                                                                                                                                               TSINGVLWTVYHGAGTRTIASPKGPVTQMYTNVDKDLVGWQAPQGSRSLTPCTCGSSDLY 120
                                                                                                                                                                                                                                                                                                                                                                                     TCINGVCWTVYHGAGTRTIASPKGPVIQMYTNVDKDLVGWPAPQGSRSLTPCTCGSSDLY 120
                                                                                                                                                                                                                                                                                                                                                                                                           180
                                                                                                                                                                                                                                                                                                                                                                                                                      9
                               Modified hepatitis C virus (HCV) NS3 protease comprising at least 1 substitution of a hydrophobic alpha-helix 0 amino acid to a hydrophilic amino acid, useful for screening inhibitors that may treat hepatitis {\sf C}
                                                                                                                                                                                                                                                                                                                                 1 MKKKGSVVIVGRINLSGDTAYAQQTRGEQGCQKTSHTGRDKNQVEGEVQIVSTATQTFLA
                                                                                                                                                                                                                                                                                                                                                                                                           LVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVSTRGVAK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Hepatitis; NS3 protease; viral replication; chronic liver disease;
                                                                                                                                                                                                                                                                                                ö
                                                                                                                                                                                                                                                                          Length 197;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Goldfarb V;
                                                                                                                                                                                                                                                                        95.4%; Score 973; DB 21;
95.4%; Pred. No. 2.6e-93;
11ve 2; Mismatches 7;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Hepatitis C virus NS4A-NS3 fusion protease #8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                liver failure; liver cancer; mutant; mutein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Zhang Y,
                                                                                       Claim 23; Fig 13; 66pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               AAB15226 standard; protein; 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                     AVDFIPVESLETTMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                AVDFIPVESLETTMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           (BRIM ) BRISTOL-MYERS SQUIBB
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Weinheimer S,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 06-JAN-2000; 2000WO-US00345
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      99US-0115271
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         19-DEC-2000 (first entry)
                                                                                                                                                                                                                                                                                    Local Similarity 95.4 les 188; Conservative
 2000-465976/40.
                                                                                                                                                                                                                                                    197 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Hepatitis C virus
           N-PSDB; AAA73330
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      WO200040707-A1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Wittekind M,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      08-JAN-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Synthetic
                                                                                                                                                                                                                                                      Sequence
                                                                                                                                                                                                                                                                                                                                                               61
                                                                                                                                                                                                                                                                                                                                                                                                         121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                          181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     AAB15226;
                                                                                                                                                                                                                                                                           Query Match
                                                                                                                                                                                                                                                                                                 Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         9
                                                                                                                                                                                                                                                                                                                                                                                                                             qq
                                                                                                                                                                                                                                                                                                                                                                õ
                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                           ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
```

```
ö
                                                                                                                                                                                                                                                                        The present sequence is a mutated version of a fusion protein created using the Hepatitis C virus (HCV) NS3 and NS4A protease enzymes. These proteins are both essential for the replication of the virus, acting to cleave its replicative proteins from the polyprotein produced from the HCV genome. Inhibitors of the two proteins should be effective as antiviral treatments of HCV infection. This is useful as HCV can lead to chronic liver disease such as cirrhosis, liver failure and liver cancer. The present invention concerns a number of NS3 mutants and NS3-NS4A well as enabling structural studies of the protease and proteins which can be used to identify inhibitors of this type, as well as enabling structural studies of the protease and protease. This sequence contains the alpha-helix0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      121 LVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVSTRGVAK 180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               9
                                                                              Modified hepatitis C virus (HCV) NS3 protease comprising at least 1 substitution of a hydrophobic alpha-helix {\bf 0} amino acid to a hydrophilic amino acid, useful for screening inhibitors that may treat hepatitis {\bf C}
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         61 TSINGVLWTVYHGAGTRTIASPKGPVTQMYTNVDKDLVGWQAPQGSRSLTPCTCGSSDLY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1 MKKKGSVVIVGRINLSGDTAYAQQTRGEQGCQKTSHTGRDKNQVEGEVQIVSTATQTFLA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Hepatitis; NS3 protease; viral replication; chronic liver disease;
liver failure; liver cancer; mutant; mutein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ó:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Length 197;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    11; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           92.7%; Score 946; DB 21; 94.4%; Pred. No. 1.7e-90;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Hepatitis C virus NS4A-NS3 fusion protease #2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 0; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Ź
                                                                                                                                                                                                                             Example 5; Fig 18; 66pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     AAB15220 standard; protein; 195
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (BRIM ) BRISTOL-MYERS SQUIBB CO.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   99US-0115271
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                06-JAN-2000; 2000WO-US00345
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      19-DEC-2000 (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Similarity 94.4
36; Conservative
WPI; 2000-465976/40
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                wild-type sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    197 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Hepatitis C virus.
Synthetic.
                           N-PSDB; AAA73335
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             W0200040707-A1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          08-JAN-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Ma.
Local 5.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       13-JUL-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  AAB15220;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Query Match
```

us-09-965-594-20.rag

```
28-JUL-1998;
28-NOV-1997;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              W09928482-A2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          24 - NOV - 1998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   10-JUN-1999
                                                                                                                                                                                                                                                               complexes.
                                                                                                                                                                                                                                                                                                                                                                                                                        29
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 179
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            AAY24943;
                                                                                                                                                                                                                                                                                       Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   RESULT 9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             AAY 24943
  ò
                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                       q
                                                                                                                                                                                                                                                                                                                                                                                                                                               ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                    셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ---
                                                                                                                                  The present sequence is a mutated version of a fusion protein created using the Hepailis C virus (HCV) NG3 and NS4A protease enzymes. These proteins are both essential for the replication of the virus, acting to cleave its replicative proteins from the polyprotein produced from the HCV genome. Inhibitors of the two proteins should be effective as antiviral treatments of the two proteins should be effective as antiviral treatments of HCV infection. This is useful as HCV can lead to chronic liver disease such as cirrhosis, liver failure and liver cancer. The present invention concerns a number of NG3 mutants and NS3-NS4A fusion proteins which can be used to identify inhibitors of this type, as well as enabling structural studies of the protease and protease: Inhibitor complexes. This sequence contains the alpha-helix0-1
                                                                                                                                                                                                                                                                                                                                                                                                  61 TSINGVLWTVYHGAGTRTIASPKGPVTQMYINVDKDLVGWQAPQGSRSLTPCTCGSSDLY 120
                                                                                                                                                                                                                                                                                                                                                                                                                                             LVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVSTRGVAK 180
                                                                                                                                                                                                                                                                                                                                                                                                                                                         119 LVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVCTRGVAK 178
                                                                                                                                                                                                                                                                                                                                                       9
                                                                                                                                                                                                                                                                                                                                                                   Modified hepatitis C virus (HCV) NS3 protease comprising at least 1 substitution of a hydrophobic alpha-helix 0 amino acid to a hydrophilic amino acid, useful for screening inhibitors that may treat hepatitis C
                                                                                                                                                                                                                                                                                                                                                      1 MKKKGSVVIVGRINLSGDTAYAQQTRGEQGCQKTSHTGRDKNQVEGEVQIVSTATQTFLA
                                                                                                                                                                                                                                                                                                                                                                                                               Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               #epatitis; NS3 protease; viral replication; chronic liver disease;
                                                                                                                                                                                                                                                                                                                                 ;;
                                                                                                                                                                                                                                                                                                          Length 195
                                                                                                                                                                                                                                                                                                                                Indels
 Goldfarb V:
                                                                                                                                                                                                                                                                                                         92.1%; Score 939; DB 21;
93.4%; Pred. No. 9.2e-90;
tive 3; Mismatches 8;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Hepatitis C virus NS4A-NS3 fusion protease #1.
 Zhang Y,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ż
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          (BRIM ) BRISTOL-MYERS SQUIBB CO.
                                                                                                               Fig 12; 66pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       AAB15212 standard; protein; 195
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      AVDFIPVESLETIMRSP 195
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           AVDFIPVESLETTMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          liver failure; liver cancer.
Wittekind M. Weinheimer S.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              36-JAN-2000; 2000WO-US00345
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  (first entry)
                                                                                                                                                                                                                                                                                                                      Best Local Similarity 93.4 Matches 184; Conservative
                      WPI; 2000-465976/40.
N-PSDB; AAA73329.
                                                                                                                                                                                                                                                                                    195 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Hepatitis C virus.
Synthetic.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                WO200040707-A1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   08-JAN-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  19-DEC-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       13-JUL-2000
                                                                                                                Claim 23;
                                                                                                                                                                                                                                                                                                                                                                                                                                              121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            AAB15212;
                                                                                                                                                                                                                                                                                    Sequence
                                                                                                                                                                                                                                                                                                           Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  RESULT 8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             AAB15212
 ò
                                                                                                                                                                                                                                                                                                                                                                           셤
                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                   ద
                                                                                                                                                                                                                                                                                                                                                                                                                                             õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                               셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
```

```
121 LVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVSTRGVAK 180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        61 TSINGVLWTVYHGAGTRTIASPRGPVTQMYTNVDRDLVGWQAPQGSRSLTPCTCGSSDLY 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            The present sequence is a fusion protein created using the Hepatitis C virus (HCV) NS3 and NS4A protease enzymes. These proteins are both essential for the replication of the virus, acting to cleave its replicative proteins from the polyprotein produced from the HCV genome. Inhibitors of the two proteins should be effective as antivixal tiveratments of HCV infection. This is useful as HCV can lead to chronic liver disease such as cirrhosis, liver failure and liver cancer. The present invention concerns a number of NS3 mutants and NS3-NS4A fusion proteins which can be used to identify inhibitors of this type, as well as enabling structural studies of the protease and protease:Inhibitor
                                                                                                                                                  Modified hepatitis C virus (HCV) NS3 protease comprising at least 1 substitution of a hydrophobic alpha-helix 0 amino acid to a hydrophilic amino acid, useful for screening inhibitors that may treat hepatitis C
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1 MKKKGSVVIVGRINLSGDTAYAQQTRGEQGCQKTSHTGRDKNQVEGEVQIVSTATQTFLA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   HCV; hepatitis C virus; single chain recombinant complex; linker; NS3 protease; NS4A cofactor; NS4A-NS3 complex; serine protease; hydrophobic domain; covalent complex; detection; inhibitor.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   .;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Length 195
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Indels
      Goldfarb V;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Score 912; DB 21;
Pred. No. 6.1e-87;
1; Mismatches 12;
   Zhang Y,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              AAY24943 standard; Protein; 665 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          HCV NS4A-NS3 complex SEQ ID NO:14.
                                                                                                                                                                                                                                                                                                           Example 2; Fig 10; 66pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     181 AVDFIPVESLETIMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Query Match 89.4%;
Best Local Similarity 92.4%;
Matches 182; Conservative 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        98US-0094331.
Wittekind M, Weinheimer S,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  98WO-US24528
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             07-SEP-1999 (first entry)
                                                            2000-465976/40.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Query Match
Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          195 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Hepatitis C virus.
Synthetic.
                                                                                            N-PSDB; AAA73328
```

```
(SCHE ) SCHERING CORP.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     RESULT 11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               AAY24942
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                q
                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                              q
                                                                                                                                                                                                                                                                                                                                                                                                                          a
                                                                                                                                                                                                                                                                                                                                                                                                                                                 õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                   q
                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                    201
                                                                                                                                                                                                                                                                                                                                                                                122 VTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVSTRGVAKA 181
                                                                                                                 The present invention describes a covalent hepatitis C virus (HCV) NS4A-NS3 complex comprising a central hydrophobic domain of native HCV NS4A peptide, a linker, and an HCV NS3 earlie protease domain, where the hydrophobic domain of native HCV NS4A peptide is tethered by the linker to the amino terminus of the HCV NS4A peptide is tethered by the linker to the amino terminus of the HCV NS3 protease domain. The present complex. The covalent NS4A NS3 complexes are useful for structural determination and determination of mode of binding of HCV inhibitors by NMR spectroscopy. They can also be used for detecting inhibitors of the protease activity, the helicase activity and the ATPase activity of NS3. The covalent NS4A-NS3 complexes are more soluble, stable and active than the non-covalent protease-peptide complexes previously available.
                                                                                                                                                                                                                                                                                                                                                                      SINGVLWTVYHGAGTRILASPKGPVTQMYTNVDKDLVGWQAPQGSRSLTPCTCGSSDLYL 121
                                                                                                                                                                                                                                                                                                                                                81
                                                                                                                                                                                                                                                                                                                                        22 GSVVIVGRIILSGSGSITAYSQQTRGLLGCKKTSLTGRDKNQVEGEVQVYSTAIQSFLAT
                                                                                                                                                                                                                                                                                                                                                                                                                            5 GSVVIVGRINLSGD -- - TAYAQQTRGEQGCQKTSHTGRDKNQVEGEVQIVSTATQTFLAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        HCV; hepatitis C virus; single chain recombinant complex; linker;
NS3 protease; NS4A cofactor; NS4A-NS3 complex; serine protease;
hydrophobic domain; covalent complex; detection; inhibitor.
                                                                                                                                                                                                                                                                                                       ÷,
                                                                                                                                                                                                                                                                                   DB 20; Length 665;
                                                                                                                                                                                                                                                                                                       Indels
                                                                                                                                                                                                                                                                                 Match 86.8%; Score 885.5; DB 20; Local Similarity 85.7%; Pred. No. 2e-83; les 168; Conservative 15; Mismatches 10;
                                ż
                                 Yao
                                                                          New hepatitis C virus covalent complexes
                                                                                                Claim 6; Page 90-92; 211pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   AAY24947 standard; Protein; 665 AA
                                 Weber PC,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   HCV NS4A-NS3 complex SEQ ID NO:18
                                                                                                                                                                                                                                                                                                                                                                                                                                                          VDFIPVESLETTMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                     VDFVPVESMETTMRSP 217
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 98WO-US24528
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      98US-0094331
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 97US-0067315
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              (first entry)
                                                                                                                                                                                                                                                             665 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Hepatitis C virus Synthetic.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     WO9928482-A2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 24-NOV-1998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               28-NOV-1997;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              07-SEP-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    28-JUL-1998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           10-JUN-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         AAY24947;
                                                                                                                                                                                                                                                                                                                                                                                                                                                          182
                                                                                                                                                                                                                                                                                                                                                                                                                                                                             202
                                                                                                                                                                                                                                                             Sequence
                                                                                                                                                                                                                                                                                   Query Match
                                                                                                                                                                                                                                                                                               Best Loca
Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                               셤
                                                                                                                                                                                                                                                                                                                                                                     à
                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
```

```
62 SINGVLWTVYHGAGTRIIASPKGPVTQMYTNVDKDLVGWQAPQGSRSLTPCTCGSSDLYL 121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    82 CVNGVCWTVYHGAGSKTLAGPKGPITQMYINVDQDLVGWQAPPGARSLTPCTCGSSDLYL 141
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      122 VTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVSTRGVAKA 181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              The present invention describes a covalent hepatitis C virus (HCV) NS4A-NS3 complex comprising a central hydrophobic domain of native HCV NS4A peptide, and an HCV NS3 serine protease domain, where the hydrophobic domain of native HCV NS4A peptide is tethered by the linker to the amino terminus of the HCV NS3 protease domain. The present sequence represents a specifically claimed example of the above complex. The covalent NS4A-NS3 complexes are useful for structural determination and determination of mode of binding of HCV inhibitors by NMR spectroscopy. They can also be used for detecting inhibitors of the protease activity, the helicase activity and the Arpase activity of NS3. The covalent NS4A-NS3 complexes are more soluble, stable and active than the non-covalent protease-peptide complexes previously available.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  S GSVVIVGRINLSGD -- - TAYAQQTRGEQGCQRTSHTGRDKNQVEGEVQIVSTATQTFLAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          22 GSVVIVGRIILSGSGSITAYSQQTRGLLGCKKTSLTGRDKNQVEGEVQVVSTATQSFLAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   HCV; hepatitis C virus; single chain recombinant complex; linker;
NS3 protease; NS4A cofactor; NS4A-NS3 complex; serine protease;
hydrophobic domain; covalent complex; detection; inhibitor.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Length 665;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Query Match 86.5%; score 882.5; DB 20; Length Best Local Similarity 85.2%; Pred. No. 4.1e-83; Matches 167; Conservative 16; Mismatches 10; Indels
                                                                     Yao N;
                                                                                                                                                                                                     New hepatitis C virus covalent complexes
                                                                                                                                                                                                                                                                   Claim 6; Page 100-102; 211pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Ź
                                                                     Weber PC,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    HCV NS4A-NS3 complex SEQ ID NO:13.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        AAY24942 standard; Protein; 665
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      VDFVPVESMETTMRSP 217
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 VDFIPVESLETTMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              98US-0094331
97US-0067315
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               98WO-US24528
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               (first entry)
                                                                     Tarem1 SS,
(SCHE ) SCHERING CORP.
                                                                                                                                   WPI; 1999-385385/32.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         665 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    W09928482-A2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               24-NOV-1998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              28-JUL-1998;
28-NOV-1997;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               07-SEP-1999
                                                              Malcolm BA,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          10-JUN-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Synthetic.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             AAY24942;
```

YAO N;

Taremi SS,

```
Malcolm BA,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                        쉱
                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                 d
                                                                                                                                                                                                                                                                                                                                                                                                                                                         á
                                                                                                                                                                                                                                                                                                                                                                                       122 VTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVSTRGVAKA 181
                                                                                                                                                                                                                                                                                                                                                                                                                                  142 VTRHADVIPVRRRGDSRGSLLSPRPVSYLKGSSGGPLLCPSGHAVGIFRAAVCTRGVAKA 201
                                                                                                   The present invention describes a covalent hepatitis C virus (HCV) NS4A-NS3 complex comprising a central hydrophobic domain of native HCV NS4A-NS3 complex. and an HCV NS3 earlie protease domain, where the hydrophobic domain of native HCV NS4A peptide is tethered by the linker to the amino terminus of the HCV NS3A protease domain. The present sequence represents a specifically claimed example of the above complex. The covalent NS4A 3C complexes are useful for structural determination and determination of mode of binding of HCV inhibitors by NMR spectroscopy. They can also be used for detecting inhibitors of the protease activity, the helicase activity and the ATPase activity of NS3. The covalent NS4A-NS3 complexes are more soluble, stable and active than the non-covalent protease-peptide complexes previously available.
                                                                                                                                                                                                                                                                                                                             5 GSVVIVGRINLSGD---TAYAQQTRGEQGCQKTSHTGRDKNQVEGEVQIVSTATQTFLAT
                                                                                                                                                                                                                                                                                                                                         22 GSVVIVGRIILSGSGSITARSQQTRGLLGCIKTSLTGRDKNQVEGEVQVVSTATQSFLAT
                                                                                                                                                                                                                                                                                                                                                                           62 SINGVLWTVYHGAGIRTIASPKGPVTQMYTNVDKDLVGWQAPQGSRSLTPCTCGSSDLYL
                                                                                                                                                                                                                                                                                                       3; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           HCV; hepatitis C virus; single chain recombinant complex; linker; NS3 protease; NS4A cofactor; NS4A-NS3 complex; serine protease; hydrophobic domain; covalent complex; detection; inhibitor.
                                                                                                                                                                                                                                                                                DB 20; Length 665;
                                                                                                                                                                                                                                                                                                      11; Indels
                                                                                                                                                                                                                                                                             86.4%; Score 881.5; DB 20
85.7%; Pred. No. 5.2e-83;
tive 14; Mismatches II.
            Yao N;
                                                         New hepatitis C virus covalent complexes
                                                                                Claim 6; Page 88-90; 211pp; English
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    AAY17880 standard; Protein; 216 AA.
            Weber PC,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       HCV NS4A-NS3 complex SEQ ID NO:4.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  111:1111:11111
202 VDFVPVESMETTMRSP 217
                                                                                                                                                                                                                                                                                                                                                                                                                                                                     182 VDFIPVESLETTMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   98US-0094331.
97US-0067315.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            98WO-US24528
                                                                                                                                                                                                                                                                                           Local Similarity 85.7%
Les 168; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       (SCHE ) SCHERING CORP
                                                                                                                                                                                                                                                          665 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              W09928482-A2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             24-NOV-1998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   28-JUL-1998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              28-NOV-1997;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 07-SEP-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    10-JUN-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Synthetic.
                                                                                                                                                                                                                                                           Sequence
                                                                                                                                                                                                                                                                                Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      å
                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                           ö
                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                              ద
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
```

```
62 SINGVLWTVYHGAGTRTIASPKGPVTQMYTNVDKDLVGWQAPQGSRSLTPCTCGSSDLYL 121
                                                                                                                                                                                           NS4A-NS3 complex comprising a central hydrophobic domain of native HCV NS4A peptide, a linker, and an HCV NS3 serine protease domain, where the hydrophobic domain of native HCV NS4A peptide is tethered by the linker to the amino terminus of the HCV NS3 protease domain. The present sequence represents a specifically claimed example of the above complex. The covalent NS4A-NS3 complexes are useful for structural determination and determination of mode of binding of HCV inhibitors by NMR spectroscopy. They can also be used for detecting inhibitors of the protease activity, the helicase activity and the Arpase activity of NS3. The covalent NS4A-NS3 complexes are more soluble, stable and active than
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            81
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  22 GSVVIVGRIILSGSGSITAYSQQTRGLLGCKKTSLTGRDKNQVEGEVQVVSTATQSFLAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  5 GSVVIVGRINLSGD---TAYAQQTRGEQGCQKTSHTGRDKNQVEGEVQIVSTATQTFLAT
                                                                                                                                                                        The present invention describes a covalent hepatitis C virus (HCV)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        HCV; hepatitis C virus; single chain recombinant complex; linker; NS3 protease; NS4A cofactor; NS4A-NS3 complex; serine protease; hydrophobic domain; covalent complex; detection; inhibitor.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  the non-covalent protease-peptide complexes previously available.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       33
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             86.1%; Score 878.5; DB 20; Length 216; 85.6%; Pred. No. 2.2e-83;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       10;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       15; Mismatches
                                                       New hepatitis C virus covalent complexes
                                                                                                              Claim 6; Page 76-77; 211pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              AAY24946 standard; Protein; 665 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       HCV NS4A-NS3 complex SEQ ID NO:17.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       VDFIPVESLETTMRS 196
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  98WO-US24528
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     98US-0094331
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             07-SEP-1999 (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Matches 167; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              (SCHE ) SCHERING CORP.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Query Match
Best Local Similarity
WPI; 1999-385385/32.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Seguence 216 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Hepatitis C virus.
Synthetic.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           24-NOV-1998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      WO9928482-A2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     28-JUL-1998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    28-NOV-1997;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                10-JUN-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     AAY24946;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          RESULT 13
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         AAY 24946
```

Weber PC, Yao N;

Taremi SS,

Malcolm BA,

œ

```
g
                                                                                                                                                                                                                                                                                                                                     õ
                                                                                                                                                                                                                                                                                                                                                         a
                                                                                                                                                                                                                                                                                                                                                                                                      쉽
                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                            82 CVNGVCWTVYHGAGSKTLAGPKGPITQMYTNVDQDLVGWQAPPGARSLTPCTCGSSDLYL 141
                                                                                                                                                                                                                                                                                                                                                SINGVLWTVYHGAGTRIIASPKGPVTQMYINVDKDLVGWQAPQGSRSLTPCTCGSSDLYL 121
                                                                                                                                                                                                                                                                                                                                                                                            122 VTRHADVIPVRRRGDSRGSLLSPRPISYIKGSSGGPLLCPAGHAVGIFRAAVSTRGVAKA 181
                                                                                                                                                                                                                                                                                                                                                                                                         NS4A NS3 complex comprising a central hydrophobic domain of native HCV NS4A Peptide, a linker, and an HCV NS3 serine protease domain, where the hydrophobic domain of native HCV NS4A peptide is tethered by the linker to the amino terminus of the HCV NS3 protease domain. The present sequence represents a specifically claimed example of the above complex. The covalent NS4A-NS3 complexes are useful for structural determination and determination of mode of binding of HCV inhibitors by NNR spectroscopy. They can also be used for detecting inhibitors of the protease activity, the helicase activity and the ArPass activity of NS3. The covalent NS4A-NS3 complexes are more soluble, stable and active than the non-covalent protease-peptide complexes previously available.
                                                                                                                                                                                                                                                                                                     61
                                                                                                                                                                                                                                                                                                                  22 GSVVIVGRIILSGSGSITAYSQQTRGLLGCIKTSLTGRDKNQVEGEVQVVSTATQSFLAT 81
                                                                                                                                                                                                                                                                                                   5 GSVVIVGRINLSGD---TAYAQQTRGEQGCQKTSHTGRDKNQVEGEVQIVSTATQTFLAT
                                                                                                                                                                                                                                                                              3; Gaps
                                                                                present invention describes a covalent hepatitis C vIrus (HCV)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             HCV; hepatitis C virus; single chain recombinant complex; linker; NS3 protease; NS4A cofactor; NS4A-NS3 complex; serine protease; hydrophobic domain; covalent complex; detection; inhibitor.
                                                                                                                                                                                                                                                      86.1%; Score 878.5; DB 20; Length 665; 85.2%; Pred. No. 1.1e-82;
                                                                                                                                                                                                                                                                              11; Indels
                                                                                                                                                                                                                                                                             15; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Weber PC, Yao N;
                                  New hepatitis C virus covalent complexes
                                                        Claim 6; Page 97-99; 211pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           HCV NS4A-NS3 complex SEQ ID NO:12.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       AAY24941 standard; Protein; 665
                                                                                                                                                                                                                                                                                                                                                                                                                                         182 VDFIPVESLETTMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                               VDFVPVESMETTMRSP 217
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             98WO-0S24528
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   98US-0094331
97US-0067315
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (first entry)
                                                                                                                                                                                                                                                                             Matches 167; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Malcolm BA, Taremi SS,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (SCHE ) SCHERING CORP.
         WPI: 1999-385385/32
                                                                                                                                                                                                                                                                  Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Hepatitis C virus.
                                                                                                                                                                                                                                 665 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  07-SEP-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             24-NOV-1998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              28-NOV-1997;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                W09928482-A2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      10-JUN-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Synthetic.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              AAY24941;
                                                                                                                                                                                                                                  Sequence
                                                                                                                                                                                                                                                                                                                                                62
                                                                                                                                                                                                                                                      Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 a
                                                                                                                                                                                                                                                                                                                                                                                                                 셤
g
                                                                                                                                                                                                                                                                                                                                                 õ
                                                                                                                                                                                                                                                                                                                                                                                            à
                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
```

```
62 SINGVLWTVYHGAGTRIIASPKGPVTQMYINVDKDLVGWQAPQGSRSLTPCTCGSSDLYL 121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   122 VTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVSTRGVAKA 181
                                                                                                                                                                    The present invention describes a covalent hepatitis C virus (HCV) NS4A-NS3 complex comprising a central hydrophobic domain of native HCV NS3 serihe protease domain, where the hydrophobic domain of native HCV NS3 serihe protease domain, where the hydrophobic domain of native HCV NS3 protease domain. The present to the amino terminus of the HCV NS3 protease domain. The present sequence represents a specifically claimed example of the above complex. The covalent NS4A-NS3 complexes are useful for structural determination and determination of mode of binding of HCV inhibitors by NMR spectroscopy. They can also be used for detecting inhibitors of the protease activity, the helicase activity and the ArPass activity of NS3. The covalent NS4A-NS3 complexes are more soluble, stable and active than the non-covalent protease-peptide complexes previously available.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              81
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       5 GSVVIVGRINLSGD -- - TAYAQQTRGEQGCQKTSHTGRDKNQVEGEVQIVSTATQTFLAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           22 GSVVIVGRIILSGSGSITAYSQQTRGLLGCRITSLTGRDKNQVEGEVQVVSTATQSFLAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             HCV; hepatitis C virus; single chain recombinant complex; linker; NS3 protease; NS4A cofactor; NS4A-NS3 complex; serine protease; hydrophobic domain; covalent complex; detection; inhibitor.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   .,
.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         DB 20; Length 665;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ž
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Yao
                                                                New hepatitis C virus covalent complexes
                                                                                                                       Claim 6; Page 85-87; 211pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    AAY17884 standard; Protein; 216 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           г,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         HCV NS4A-NS3 complex SEQ ID NO:8.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Weber
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  VDFVPVESMETTMRSP 217
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 182 VDFIPVESLETIMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    98WO-0S24528
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        98US-0094331.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            07-SEP-1999 (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Query Match 86.0 Best Local Similarity 85.2 Matches 167; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Taremi SS,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (SCHE ) SCHERING CORP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              WPI; 1999-385385/32
WPI; 1999-385385/32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   665 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   W09928482-A2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           28-JUL-1998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       28-NOV-1997;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Malcolm BA,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             10-JUN-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Synthetic.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           AAY17884;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   AAY17884
```

```
The present invention describes a covalent hepatitis C virus (HCV) NS4A-NS3 complex comprising a central hydrophobic domain of native HCV NS4 aperide, a linker, and an HCV NS3 earlne protease domain, where the hydrophobic domain of native HCV NS4A peptide is tethered by the linker to the amino terminus of the HCV NS4A peptide is tethered by the linker to the amino terminus of the HCV NS3 protease domain. The present sequence represents a specifically claimed example of the above complex. The covalent NS4A NS3 complexes are useful for structural determination and determination of mode of binding of HCV inhibitors by MNR spectroscopy. They can also be used for detecting inhibitors of the protease activity, the helicase activity and the ATPase activity of NS3. The covalent NS4A NS3 complexes are more soluble, stable and active than the non-covalent protease-peptide complexes previously available.
                           New hepatitis C virus covalent complexes
                                                                                         Claim 6; Page 80-81; 211pp; English.
```

216 AA; Sequence

ij 62 SINGVLWTVYHGAGTRTIASPKGPVTQMYTNVDKDLVGWQAPQGSRSLTPCTCGSSDLYL 121 122 VTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVSTRGVAKA 181 5 GSVVIVGRINLSGD----TAYAQQTRGEQGCQKTSHTGRDKNQVEGEVQIVSTATQTFLAT 61 22 GSVVIVGRIILSGSGSITAYSQQTRGLLGCKKTSLTGRDKNQVEGEVQVVSTATQSFLAT 81 3; Gaps Query Match 85.8%; Score 875.5; DB 20; Length 216; Best Local Similarity 85.1%; Pred. No. 4.6e-83; Matches 166; Conservative 16; Mismatches 10; Indels 3; 182 VDFIPVESLETIMRS 196 |||::|||:|||||| 202 VDFVPVESMETIMRS 216 đ g õ qq ò ò õ

Search completed: August 30, 2003, 19:12:24 Job time : 44.6227 secs

g

```
GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.
```

using sw model - protein search, OM protein August 30, 2003, 19:02:22; Search time 16.2134 Seconds (without alignments) 1168.492 Million cell updates/sec ü

US-09-965-594-20 1020

1 MKKKGSVVIVGRINLSGDTA......vakavDFIPVESLETIMRSP 197 Perfect score: Sequence:

BLOSUM62 Gapop 10.0 , Gapext 0.5 Scoring table:

283308 seqs, 96168682 residues Searched: Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0 Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database

PIR\_76:\*
1: pir1:\*
2: pir2:\*
3: pir3:\*
3: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Description		genome polyprotein										genome polyprotein	polyprotein - dour	polyprotein - marm	probable aromatic	probable periplasm	hypothetical prote	•~	nitrate/nitrite se	hypothetical prote	ω,	low density lipopr	probable oxidoredu	htrA-like serine p	beta transducin-li	probable sugar upt	hypothetical prote		U	glycosyl transfera
	Ü		GNWVC3	S40770	GNWVCH	GNWVTW	A45573	GNWVIC	GNWVCJ	S18030	JC5620	GNWVJ8	JQ1303	T08841	T08839	H83144	B71284	G83612	B81104	C81911	A75335	139383	JE0372	H70847	H97199	T42045	F95973	B71360	A49448	T04533	D87603
	DB	١,	_	-	-	-	-	7	-	Н	Η,	н	٦	N	~	~	~	~	~	7	~	ď	7	7	~	~	~				
	Length		3011	3011	3011	3010	3010	3010	3010	3010	3014	3033	3033				398	716	290	290	377	452	1615	479	433	1049	394	495	764	846	322
æ	Query Match		83	83	83	82	81	80	80	7	72	99	99	7	N	œ				8.2		œ	æ	7 .	7.	7.		7	7	7.7	7.
	Score		654.5	853.5	848.5	836.5	827.5	823.5	823.5	811.5	743.5	675	673	251	245	85.5	85.5	84.5	83.5	83.5	83	83	83	80.5	79.5	79.5	78.5	78.5	78.5	78.5	78
,	Result No.		-	7	m	❖	S	Q	7	æ	6	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	56	27	28	53

ä

..

DB 1; Length 3011; 17; Indels

Query Match 83.8%; Score 854.5; DB 1 Best Local Similarity 82.8%; Pred. No. 9.7e-70; Matches 169; Conservative 9; Mismatches 17

ð

hypothetical prote	HIV-1 retropepsin	hypothetical prote	proteinase (EC 3.4	mosA protein - Rhi	probable exported	hypothetical prote	transferin-bindin	qenome polyprotein	genome polyprotein	genome polyprotein	hypothetical prote	hypothetical prote	hypothetical prote	requiatory protein	hypothetical prote
E75392	GNLJGL	S57895	PRLJHD	B53308	AH1030	A81942	F81196	GNWVY	GNWVYP	GNWVNE	T32434	G90654	G85505	E64745	BB3477
~1	7	~	~	н	~	~	~	٦	~	-	ď	~	~	~	7
477	1165	361	323	333	492	539	915	3411	3411	3414	424	452	452	452	515
7.6	7.6	7.6	7.5	7.5	7.5	7.5	7.5	7.5	7.5	7.5	7.4	7.4	7.4	7.4	7.4
78	78	77.5	76.5	76.5	76.5	76.5	97	16	92	76	75.5	75.5	75.5	75.5	75.5

```
A;Acteus: preliminary
A;Acteus: preliminary
A;Acteus: preliminary
A;Acteus: preliminary
A;Acteus: preliminary
A;Acteus: preliminary
A;Acteus: acceptance: solutes
A;Resperimental source: solutes
C;Superfamily: hepatitis C virus genome polyprotein
C;Superfamily: hepatitis C virus genome polyprotein
C;Superdamily: hepatitis C virus genome polyprotein
C;Superdamily: hepatitis C virus genome polyprotein
C;Superdamily: hepatitis C virus genome polyprotein
F;1-115/Product: capsid protein C *status predicted <RED>
F;101-115/Product: envelope protein M *status predicted <RED>
F;102-194/Product: nonstructural protein NSI *status predicted <NSI>
F;1007-165/Product: nonstructural protein NSI *status predicted <NSI>
F;1030-1213/Region: nucleotide-binding motif A (P-loop)
F;1312-1317/Region: DEXH motif
genome polyprotein - hepatitis C virus (strain HCV-1)
N;Contains: capsid protein C; envelope protein M; hepacivirin (EC 3.4.21.98) (nonstruprotein NS4s: nonstructural protein NS4s; nonstructural protein NS5s:
C;Species: hepatitis C virus
C;Date: 30-Sep-1992 #sequence_revision 30-Sep-1992 #text_change 19-Jan-2001
C;Choc, O.L.; Richman, K.H.; Han, J.H.; Berger, K.; Lee, C.; Dong, C.; Gallegos, C.; Proc. Natl. Acad. Sci U.S.A. 88, 2451-2455, 1991
A;Title: Genetic organization and diversity of the hepatitis C virus.
A;Reference number: A39166; MUID:91172826; PMID:1848704
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      F;1616-1862/Product: nonstructural protein NS4a *status predicted <N4A>
F;1863-2013/Product: nonstructural protein NS4b *status predicted <N4B>
F;2014-3011/Product: nonstructural protein NS5 *status predicted <NS5>
F;2014-3011/Product: nonstructural protein NS5 *status predicted <NS5>
F;196,209,234,305,325,417,423,430,448,476,532,540,556,576,623,645,1213,1255,2041,2077
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            A.Molecule type: mRNA
A.Molecule type: mRNA
A.Molecule type: mRNA
A.Residues: 1-3011 <cHo>
A.Cross-references: GB:M62321; NID:9329873; PIDN:AAA45676.1; PID:9329874
S.Chan, S.W.; McOmish, F.; Holmes, E.C.; Dow, B.; Peutherer, J.F.; Follett, E.; Yap,
J. Gen. Virol. 73, 1131-1141, 1992
A.Title: Analysis of a new hepatitis C virus type and its phylogenetic relationship (
A.Reference number: PO0403
A.Recession: PQ0403
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             A; Molecule type: genomic RNA
A; Residues: 1577-1633 <CHA>
A;Cross-references: DDBJ:D10128
A;Experimental source: isolates E-b16
A;Accession: PQ0404
```

us-09-965-594-20.rpr

~

```
A; Molecule type: genomic RNA
A; Residues: 1-3011 < INC>
A; Residues: 1-3011 < INC>
A; Cross-references: GB: M67463; NID: 9329737; PIDN: AAA45534.1; PID: 9329738
A; Cross-references: GB: M67463; NID: 929737; PIDN: AAA45534.1; PID: 9329738
A; Inchauspe, G.; Zebedee, S.; Lee, D.H.; Sugitani, M.; Nasoff, W.; Prince, A.M.
Proc. Natl. Acad. Sci. U.S.A. 88, 10292-10296, 1991
A; Title: Genomic structure of the human prototype strain H of hepatitis C virus: comp A; Reference number: A41546; MUID: 92052256; PMID: 1658800
                                                                                                                                                                                                                                                                                      C.Species: hepatitis C virus
A:Note: host Homo sapiens (man)
C.Spacies: bepatitis C virus
A:Note: host Homo sapiens (man)
C.Spacies: J.Dec-1992 *sequence_revision 31-Dec-1992 *text_change 19-Jan-2001
C.Spacession: A36814; A41546
B:Inchauspe, G.: Zebedee, S.; Lee, D.H.; Sugitani, M.; Nasoff, M.; Prince, A.M.
Submitted to GenBank, July 1992
A;Description: Genomatic structure of the human prototype strain H of hepatitis C virus
A;Reference number: A36814
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       A:Note: neither amino acid nor nucleotide sequence is given
C:Superfamily: hepatitis C virus genome polyprotein
C:Superfamily: hepatitis C virus genome polyprotein
C:Superfamily: hepatitis C virus genome polyprotein; hydrolase; nonstruct
C:Superfamily: hepatitis C virus genome polyprotein; hydrolase; nonstruct
C:Superfamily: hydroclain; envelope protein C *status predicted <CPC>
F:115/Product: capsid protein C *status predicted <CPN>
F:192-389/Product: major envelope protein M *status predicted <MEE>
F:300-729/Product: nonstructural protein NSI *status predicted <NSI>
F:730-1006/Product: nonstructural protein NSI *status predicted <NSI>
F:1230-1237/Region: nucleotide-binding motif A (P-loop)
F:1312-1317/Region: DEXH motif
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  genome polyprotein - hepatitis C virus (strain Taiwan)
N:Contains: capsid protein C; envelope protein M; hepacivirin (EC 3.4.21.98) (nonstru-
protein NS4a; nonstructural protein NS4b; nonstructural protein NS5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     F;1616-1862/Product: nonstructural protein NS4a *status predicted <N4A>
F;1863-2013-7Product: nonstructural protein NS4b *status predicted <NN4B>
F;2014-1901/Product: nonstructural protein NS5 *status predicted <NN5P
F;196,209,234,305,325,417,423,430,448,476,932,540,556,576,623,445,1213,1255,2041,2240
                                                                                                                                                                                                                          N'Contains: capsid protéin C; envelope protein M; hepacivirin (EC 3.4.21.98) (nonstru
protein NS4a; nonstructural protein NS4b; nonstructural protein NS5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1065 ATQTFLATCINGVCWTVYHGAGTRTIASPKGPVIQTYTNVDQDLVGWPAPQGSRSLTPCT 1124
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     54 ATGIFLATSINGVLMTVYHGAGTRIIASPKGPVTQMYTNVDKDLVGWQAPQGSRSLIPCT 113
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      114 CGSSDLYLVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAV 173
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               C.Species: hepatitis C virus
A:Note: host Homo sapiens (man)
C.Jate: 31-Dec-1992 *sequence_revision 31-Dec-1992 *text_change 19-Jan-2001
C.Jace: 31-Dec-1992 *sequence_revision 31-Dec-1992 *text_change 19-Jan-2001
C.Jacession: A40244
R:Chen, P.J.; Lin, M.H.; Tai, K.F.; Liu, P.C.; Lin, C.J.; Chen, D.S.
Virology 188, 102-113, 1992
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             3 KKGSVVIVGRIN-----LSGDTAYAQQTRGEQGCQKTSHTGRDKNQVEGEVQIVST
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              9; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Score 848.5; DB 1; Length 3011;
Pred. No. 3.5e-69;
10; Mismatches 18; Indels 9;
                                                                                                                                                                       genome polyprotein - hepatitis C virus (strain H) N:Contains: capsid protein C; envelope protein H;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                174 STRGVAKAVDFIPVESLETTMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      83.2%;
81.9%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Matches 167; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Query Match
Best Local 9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 d
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      å
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              A. Molecule type: genomic RNA
A. Residues: 1-513 < colors
A. Molecule type: genomic RNA
A. Residues: 1-513 < colors
A. Cross references: GBs 1008311; NID: 9221511; PIDN: BAA00705.1; PID: 9221512
A. Cross references: GBs 1008311; NID: 9221511; PIDN: BAA00705.1; PID: 9221512
C. Superfamily: heparitis C virus genome polyprotein
C. Superfamily: Polop: polyprotein
C
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     genome polyprotein - hepatitis C virus
N;Contains: capsid protein C; envelope protein M; hepacivirin (EC 3.4.21.98) (nonstructu
N;Contains: capsid protein C; envelope protein MS4s; nonstructural protein NS5s; nonstructural protein NS4s; nonstructural protein NS4s; nonstructural protein C; Species: hepatitis C virus
C; Species: l9-May-2000 *sequence_revision 19-May-2000 *text_change 19-Jan-2001
C; Accession: S40770; PC1285
R; Okamoto, H.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              A; Cross references: EMBL:D10749; NID:q221586; FIDN:BRA01582.1; PID:q221587
R; Okamoto, H.; Okada, S.; Sugiyama, Y.; Yotsumoto, S.; Tanaka, T.; Yoshizawa, H.; Tsuda, Jpn. J. Exp. Med. 60, 167-177, 1990
A; Title: The S'-terminal sequence of the hepatitis C virus genome.
A; Reference number: PC1284; MUID:91013116; PMID:2170712
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1005 RKGREILLGPADGMVSKGWRLLAPITAYAQQTRGLLGCIITSLTGRDKNQVEGEVQIVST 1064
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 54 ATQTFLATSINGVLWTVYHGAGTRTIASPKGPVTQMYTNVDKDLVGWQAPQGSRSLTPCT 113
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          114 CGSSDLYLVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAV 173
.005 RRGREILLGPADGMVSKGWRLLAPITAYAQQTRGLLGCIITSLTGRDKNQVEGEVQIVST 1064
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       3 KKGSVVIVGRIN----LSGDTAYAQQIRGEQGCQKTSHTGRDKNQVEGEVQIVST
                                                                                        54 ATOTFLATSINGVLWTVYHGAGTRTIASPKGPVTOMYTNVDKDLVGWQAPQGSRSLTPCT
                                                                                                                                                                                                                                                                   114 CGSSDLYLVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    F;1616-1862/Product: nonstructural protein NS4a *status predicted <N4A>F;1863-2013/Product: nonstructural protein NS4b *status predicted <N4B>F;2014-3011/Product: nonstructural protein NS5 *status predicted <NS5>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Length 3011;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    18; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     DB 1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          2e-69;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Query Match 83.7%; Score 853.5; Best Local Similarity 82.8%; Pred. No. 1.2e Matches 169; Conservative 8; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Score 853.5;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         submitted to the EWBL Data Library, March 1992
A; Reference number: $40770
A; Accession: $40770
                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1185 CTRGVAKAVDFIPVESLETTMRSP 1208
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     174 STRGVAKAVDFIPVESLETTMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                      174 STRGVAKAVDFIPVESLETTMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  A; Molecule type: genomic RNA
A; Residues: 1-3011 < OKA>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  A: Accession: PC1285
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ô
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ö
                                                                                                                                                                                                                                            ŏ
                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           q
                                                                                   ò
                                                                                                                                                         q
```

m

```
Query Match
Best Local Similarity 76.5
Matches 156; Conservative
                                                                                                                                                                                                                                                  Conservative
                                                                                                                                                                                                             Similarity
                                                                                                                                                                    Query Match
Best Local Simi
Matches 157;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      RESULT 6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                õ
                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         C; Species: hepatitis C virus
C; Date: 19-May-2000 #sequence_revision 19-May-2000 #text_change 19-Jan-2001
C; Accession: A45573
Virus Res. 23, 39-53, 1992
A; Title: Wolecular cloning of hepatitis C virus genome from a single Japanese carrier: 8
A; Reference number: A45573; MuiD:92295714; PMID:1318627
A; Accession: A45573
A; Accession: A4
                                                                                                                                                                                A; Cross-references: CB:M84754
C; Superfamily: hepatitis C virus genome polyprotein
C; Superfamily: hepatitis C virus genome polyprotein; G; Superfamily: hepatitis C virus genome protein; glycoprotein; hydrolase; nonstructura
C; Superfamily: hepatitis C virus genome protein; glycoprotein; hydrolase; nonstructural
F; 1-115/Product: capsid protein C *status predicted <CPR>F; 116-191/Product: major envelope protein B *status predicted <MEE>F; 390-729/Product: nonstructural protein NSI *status predicted <NSI>F; 300-71615/Product: nonstructural protein NSI *status predicted <NSI>F; 1300-1615/Product: hepativirin *status predicted <NSI>F; 1317/Region: nucleotide-binding motif A (P-loop)
F; 1315-1317/Region: DEXH motif
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      genome polyprotein - hepatitis C virus (strain JT)
N.Contains: capsid protein C; envelope protein M; hepacivirin (EC 3.4.21.98) (nonstructu
protein NS4a; nonstructural protein NS4b; nonstructural protein NS5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             F;1616-1862/Product: nonstructural protein NS4a #status predicted <N4A>
F;1863-2013/Product: nonstructural protein NS4b #status predicted <N4B>
F;2014-3010/Product: nonstructural protein NS5 #status predicted <NS5>
F;196,209,233,234,250,305,325,417,423,430,448,532,540,556,576,623,645,1213,1255,2041,207
A;Title: The Taiwanese hepatitis C virus genome: sequence determination and mapping A;Reference number: A40244; MUID:92230206; PMID:1314449
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1005 RRGREILLGPADSLEGRGWRLLAPITAYAQQTRGLFGCIITSLTGRDKNOVEGEVOVVST 1064
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               113
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   114 CGSSDLYLVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAV 173
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       3 KKGSVVIVGRIN-----LSGDTAYAQQTRGEQGCQKTSHTGRDKNQVEGEVQIVST
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ATQTFLATSINGVLWTVYHGAGTRTIASPKGPVTQMYTNVDKDLVGWQAPQGSRSLTPCT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Length 3010;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              DB 1:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          82.0%; Score 836.5; DB 1
78.4%; Pred. No. 4.4e-68;
Live 18; Mismatches 17
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1185 CTRGVAKAVDFVPVESMETTMRSP 1208
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       174 STRGVAKAVDFIPVESLETTMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  F;1316-1319/Region: DEXH motif
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Best Local Similarity 78.4
Matches 160; Conservative
                                                                                                                               A; Molecule type: genomic RNA A; Residues: 1-3010 <CHE>
                                                                                              A; Accession: A40244
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
```

```
A; Molecule type: genomic RNA
A; Residues: 1-3010 cTAK7
A; Residues: 1-3010 cTAK7
A; Residues: 1-3010 cTAK7
A; Cass references: EMBL:M58335; NID:g129770; PIDN:AAA72945.1; PID:g329771
C; Superfamily: hepatitis C virus genome polyprotein
C; Keywords: ATP; capsid protein; envelope protein; glycoprotein; hydrolase; nonstruc
F; 2-115/Product: capsid protein C *status predicted <EPN-
F; 116-191/Product: major envelope protein E *status predicted <HEE>
F; 192-389/Product: major envelope protein R1 *status predicted <HS1>
F; 390-729/Product: nonstructural protein NS1 *status predicted <HS2>
F; 1007-1615/Product: nonstructural protein NS2 *status predicted <HS2>
F; 1130-1217/Region: nucleotide-binding motif A (P-loop)
F; 1116-1119/Region: DEXH motif
F; 11616-1180/Product: nonstructural protein NS4* *status predicted <HAP>
F; 1863-2013/Product: nonstructural protein NS4* *status predicted <HAP>
F; 1863-2013/Product: nonstructural protein NS5* *status predicted <HS2>
F; 2014-3010/Product: nonstructural protein NS5* *status predicted <HS2>
F; 2014-3010/FR3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       genome polyprotein - hepatitis C virus
N;Contains: capsid protein C; envelope protein M; hepacivirin (EC 3.4.21.98) (nonstr
protein N344; nonstructural protein N34b; nonstructural protein N35
C;Species: hepatitis C virus
C;Species: 31-Mar-1992 *text_change 19-Jan-2001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     C;Accession: A38465
R:Takamizawa, A.; Mozi, C.; Fuke, I.; Manabe, S.; Murakami, S.; Fujita, J.; Onishi, J. Yirol. 65, 1105-1113, 1991
A;Title: Structure and organization of the hepatitis C virus genome isolated from hu A;Reference number: A38465; MUID:91140698; PMID:1847440
A;Reference number: A38465
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       F;196,209,234,250,305,325,417,423,430,448,532,540,556,576,623,645,1213,1255,2041,207
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1124
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              54 ATQTFLATSINGVLWTVYHGAGTRTIASPKGPVTQMYTNVDKDLVGWQAPQGSRSLTPCT 113
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       3 KKGSVVIVGRIN------LSGDTAYAQQTRGEQGCQKTSHTGRDKNQVEGEVQIVST
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  114 CGSSDLYLVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1005 REGKEILLGPADSLEGRGLRLLAPITAYSQQTRGLLGCITTSLTGRDKNQVEGEVQVVST
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1065 ATQSFLATCVNGVCWTVYHGAGSKTLAAPKGPITQMYTNVDQDLVGWPKPPGARSLTPCT
                                                                                                                                                                                                                                                                                                                                                3 KKGSVVIVGRIN-----LSGDTAYAQQTRGEQGCQKTSHTGRDKNQVEGEVQIVST
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Gaps
                                                                                                                                                                                                                                                                   Gaps
F;1616-1862/Product: nonstructural protein NS4a *status predicted <N4A>F;1863-2013/Product: nonstructural protein NS4b *status predicted <N4B>F;2014-3010/Product: nonstructural protein NS5 *status predicted <NS5>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ٠<u>.</u>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Length 3010;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Indels
                                                                                                                                                                     81.1%; Score 827.5; DB 1; 77.0%; Pred. No. 2.9e-67;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 DB 1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            80.7%; Score 823.5; DB 1
76.5%; Pred. No. 6.8e-67;
tive 21; Mismatches 18
                                                                                                                                                                                                                                                      20; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 STRGVAKAVDFIPVESLETIMRSP 197
```

ă

ô g

```
A.Molecule type: genomic RNA
A.Residues: 1-547, 'T',549-621,'V',623-624,'S',626-652,'DL',655-761,'T',763-782 A.Cross-references: EMBL:X61591
A.Note: this sequence is inconsistent with the nucleotide translation
A.Note: this sequence is inconsistent with the nucleotide translation
A.Note: the authors translated the codon AGG for residue 43 as Pro, TGG for residue:
as TIP, and TTC for residue 77 as Ser
A.Note: sequence extracted from NCBI backbone (NCBIN:121747, NCBIP:121748)
C.Superfamily: hepatitis C virus genome polyprotein
C.Superfamily: hepatitis C virus genome predicted <CEPA-
F.115/Product: envelope protein MS1 *status predicted <MS2-
F.130-123/Product: nonstructural protein NS1 *status predicted <NS2-
F.130-123/Product: nonstructural protein NS1 *status predicted <NS2-
F.1310-133/Region: nucleotide-binding motif B
F.1312-1337/Region: nucleotide-binding motif B
F.1316-1862/Product: nonstructural protein NS4 *status predicted <N4A-
F.1816-1862/Product: nonstructural protein NS4 *status predicted <NAA-
F.1816-1801/Product: nonstructural protein NS4 *status predicted <NAS-
F.1816-209,234,250,305,417,423,448,532,540,556,576,623,645/Binding site: carbohydrate
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    genome polyprotein - hepatitis C virus (isolate EUH1480)
N;Contains: capsid protein C; envelope protein M; hepacivirin (EC 3.4.21.98) (nonstruprotein NS4a; nonstructural protein NS4b; nonstructural protein NS5 C;Species: hepatitis C virus
C;bate: 19-May-2000 #sequence_revision 19-May-2000 #text_change 19-Jan-2001
C;Accession: JC5620
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Richamberlain, R.W.; Adams, N.J.; Taylor, L.A.; Simmonds, P.; Elliott, R.M.
Blochem. Biophys. Res. Commun. 236, 44-49, 1997
A; Title: The complete coding sequence of hepatitis C virus genotype 5a, the predomins
A; Reference number: JC5620; MUID:97366593; PMID:9223423
                                                                                                                                                                                                                                                                                                                                                                                                                                                               A:Experimental source: isolate JK1 from an individual
R:Honda, M.; Kaneko, S.; Unoura, M.; Kobayashi, K.; Murakami, S.
Arch. Virol. 128, 163-169, 1993
A;Title: Sequence analysis of putative structural regions of hepatitis C virus isolat A;Reference number: A48332; MUID:93119270; PMID:8380322
A;Accession: S33570
                                                                                                                                                                     A; Description: A whole genome of hepatitis C virus cDNA was isolated from a single A; Reference number: $18028
A; Accession: $18030
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1125 YGSSDLYLVTRHADVIPVRRKGDSRGSLLSPRPVSYLKGSSGGPLLCPSGHAVGIFRAAV 1184
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  114 CGSSDLYLVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAV 173
                                           C;Date: 19-May-2000 *sequence_revision 19-May-2000 *text_change 23-Mar-2001 C;Cacession: SIBOJO; S33570; A48329; S18029 R;Honda, M.; Kaneko, S.; Masashi, U.; Kobayashi, K.; Murakami, S. submitted to the EMBL Data Library, September 1991
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         3 KKGSVVIVGRIN-----LSGDTAYAQQTRGEQGCQKTSHTGRDKNQVEGEVQIVST
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     54 ATQTFLATSINGVLWTVYHGAGTRTIASPKGPVTQMYTNVDKDLVGWQAPQGSRSLTPCT
                                                                                                                                                                                                                                                                                                                                          A;Molecule type: genomic RNA
A;Residues: 1-3010 <RON>
A;Cross-references: EMBL:X61596; NID:g59478; PIDN:CAA43793.1; PID:g59479
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    79.6%; Score 811.5; DB 1; Length 3010; ilarity 76.0%; Pred. No. 8.7e-66; Conservative 20; Mismatches 20; Indels 9;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       174 STRGVAKAVDFIPVESLETTMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Local Similarity
nes 155; Conserv
    A; Variety: isolate JK1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Best Loca
Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               JC5620
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      A. Molecule type: genomic RNA
A. Residues: 2650-2707 < kA2>
A. Molecule type: genomic RNA
A. Residues: 2650-2707 < kA2>
C. Comment: The cleavage sites of this polyprotein by the cleavage sites of this polyprotein
C. Superfamily: hepatitis C virus genome polyprotein
F. 2-115/Product: envelope protein C *status predicted < REDN>
F. 1912-289/Product: envelope protein NS1 *status predicted < NS2>
F. 1907-165/Product: nonstructural protein NS2 *status predicted < NS2>
F. 1230-1237/Region: nucleotide-binding motif A (P-loop)
F. 1312-1319/Region: nucleotide-binding motif B
F. 1316-1319/Region: DEXH motif
F. 1661-2013/Product: nonstructural protein NS4 *status predicted < N4A>
F. 2014-3010/Product: nonstructural protein NS5 *status predicted < NS5>
F. 1965-2013/Product: nonstructural protein NS5 *status predicted < NS5>
F. 1966-2034, 250, 305, 325, 417, 423, 430, 448, 532, 556, 576, 623, 645, 1213, 1255, 2041, 2077, 2240, 2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Rivato, N.; Hijkata, M.; Ootsuyama, Y.; Nakagawa, M.; Ohkoshi, S.; Sugimura, T.; Shimop Proc. Natl. Acad. Sci. U.S.A. 87, 9524-9528, 1990
Aritle: Molecular cloning of the human hepatitis C virus genome from Japanese patients A; Reference number: A39253; MUD:91088550; PMID:2175903
A; Accession: A39253, MUD:91088550; PMID:2175903
A; Accession: A39253
A; Mulcular type: genomic RNA
A; Residues: 1-3010 (KAr>
A; Coss-references: GB:D90208; NID:g221610; PIDN:BAA14233.1; PID:g221611
B; Kato, N.; Ohkoshi, S.; Shimocohno, K.
Proc. Jpn. Acad. 65B, 219-223, 1989
A; Title: Japanese isolates of the non-A, non-B hepatitis viral genome show sequence variance number: PS0085
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 genome polyprotein - hepatitis C virus (isolate JK1)
N;Contains: capsid protein C; envelope protein M; hepacivirin (EC 3.4.21.98) (nonstructu
protein NS4a; nonstructural protein NS4b; nonstructural protein NS5
C;Species: hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                   nonstructura
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1005 RRGKEILLGPADSFGEQGWRLLAPITAYSQQTRGLLGC11TSLTGRDKNQVDGEVQVLST 1064
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       54 ATQTFLATSINGVLWTVYHGAGTRTIASPKGPVTQMYTNVDKDLYGWQAPQGSRSLTPCT 113
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     114 CGSSDLYLVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAV 173
                                                                                                                                                                                                                                                                                      genome polyprotein - hepatitis C virus (strain J)
N;Contains: capsid protein C; envelope protein M; major envelope protein E;
protein NS4a; nonstructural protein NS4b; nonstructural protein NS5
                                                                                                                                                                                                                                                                                                                                                                                                                    C.Species: hepatitis C virus
C.Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 19-Jan-2001
C.Accession: A39253; PS0086
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               3 KKGSVVIVGRINLSGD-----TAYAQOTRGEQGCOKTSHTGRDKNOVEGEVOIVST
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          80.7%; Score 823.5; DB 1; Length 3010; 75.5%; Pred. No. 6.8e-67; Live 23; Mismatches 18; Indels 9;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 174 STRGVAKAVDFIPVESLETTMRSP 197
174 STRGVAKAVDFIPVESLETTMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Query Match
Best Local Similarity 75.55
Matches 154; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    A; Accession: PS0086
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         $18030
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         엄
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               셤
```

ŏ

ò

ò

```
Query Match 66.2%;
Best Local Similarity 69.8%;
Matches 125; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Ouery Match 66.0%
Best Local Similarity 68.7%
Matches 123; Conservative
                                                                                          <KAT>
                                        A; Molecule type: mRNA
A; Residues: 2678-2729
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           RESULT 11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ô
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ŝ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         a
A. Residues: 1-3014 CCHA>
A. Residues: 1-3014 CCHA>
A. Residues: 1-3014 CCHA>
A. Residues: 1-3014 CCHA>
A. Experimental source: genotype 5a, which predominates in South Africa
A. Experimental source: genotype 5a, which predominates in South Africa
A. Note: the translation of the nucleotide sequence is not complete in this paper
C. Superfamily: hepatitis C virus genome polyprotein
C. Superfamily: hepatitis C virus genome polyprotein
C. Reywords: ATP: glycoprotein: hydrolase: nucleotide binding; P-loop; polyprotein; C. Reywords: ATP: glycoprotein: hydrolase: nucleotide brains predicted CPPA>
F. 192-189/Product: major envelope protein R status predicted CME>
F. 384-408/Region: hypervariable *status predicted cNSI>
F. 391-1007/Product: nonstructural protein NSI *status predicted cNSI>
F. 1123-1128CAPEGION: nucleotide-binding motif A (P-loop)
F. 1317-1320/Region: nucleotide-binding motif A (P-loop)
F. 1317-1320/Region: DEXH motif
F. 1864-2014/Product: nonstructural protein NS4* *status predicted cN4>
F. 2015-3014/Product: nonstructural protein NS4* *status predicted cN5>
F. 2015-3014/Product: nonstructural protein NS5 *status predicted cN5>
F. 2015-3014/Froduct: nonstructural protein NS5 *status predicted cN5>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     C. Species: heparitis C virus
C. Species: heparitis C virus
C. Date: 31-Dec-1992 #sequence_revision 31-Dec-1992 #text_change 19-Jan-2001
C. Date: 31-Dec-1992 #sequence_revision 31-Dec-1992 #text_change 19-Jan-2001
C. Accession: Ado250: p00397: p00559
R. Accession: Ado250: p00397: p00559
R. Accession: Ado250: p00397: p00559
A. Feference number: Ado250: mUID:92230232; pMID:1314459
A. Accession: A40250
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  A:Experimental source: isolate E-bl2
R;Kato, N.; Octsuyama, Y.; Ohkoshi, S.; Nakazawa, T.; Mori, S.; Hijikata, M.; Shimotohno
Blochem Biophys. Res. Commun. 181, 279-285, 1991
A;Title: Distribution of plural HCV types in Japan.
A;Reference number: PQ0554; MUID:92068204; PMID:1720309
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       genome polyprotein - hepatitis C virus (strain HC-J8)
N:Contains: capsid protein C; envelope protein M; hepacivirin (EC 3.4.21.98) (nonstructu
protein NS4a; nonstructural protein NS4b; nonstructural protein NS5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             A: Molecule type: genomic RNA
A: Residues: 1-3033 < COXA>
A: Residues: 1-3033 < COXA>
A: COXA>
A: CTOSS-TEFERENCES: GB: CD102021; NID: g221608; PIDN: BAA01761.1; PID: g221609
A: CTOSS-TEFERENCES: GB: D10988; GB: D01221; NID: g221608; PIDN: BAA01761.1; PID: g221609
B: Chan, S: W.: McOmish, F.; Holmes, E.C.; Dow, B.; Peutherer, J.F.; Follett, E.; Yap, P.I.
J. Gen. Virol. 73, 1131-1141, 1992
A: Title: Analysis of a new hepatitis C virus type and its phylogenetic relationship to A: Reference number: PQ0393; MUID: 92268871; PMID: 1316939
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1126 CGSADLYLVTRHADVIPARRRGDTRASLLSPRPISYLKGSSGGPIMCPSGHVVGVFRAAV 1185
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ...
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1006 RRGREIFLGPADDIKTSGWRLLAPITAYAQQTRGVLGAIVLSLTGRDKNEAEGEVQFLST 1065
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        54 ATQTFLATSINGVLWTVYHGAGTRTIASPKGPVTQMYTNVDKDLVGWQAPQGSRSLTPCT 113
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 114 CGSSDLYLVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAV 173
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   3 KKGSVVIVGRIN-----LSGDTAYAQQTRGEQGCQKTSHTGRDKNQVEGEVQIVST 53
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Query Match 72.9%; Score 743.5; DB 1; Length 3014; Best Local Similarity 68.6%; Pred. No. 1.5e-59; Matches 140; Conservative 25; Mismatches 30; Indels 9;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  174 STRGVAKAVDFIPVESLETIMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          A; Cross-references: DDBJ:D10134
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          A: Molecule type: genomic RNA A: Residues: 2678-2754 <CHA>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             A; Accession: PQ0397
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           q
```

```
A/Tross references: GB:D10562; GB:D90518; NID:q221523; PIDN:BAA01418.1; PID:q221524
C; Superfamily: hepatitis C virus genome polyprotein
F:116-191/Product: capsid protein C *status predicted <BPA>
F:116-191/Product: envelope protein E *status predicted <BPA>
F:136-191/Product: monstructural protein NSI *status predicted <NSI>
F:134-1010/Product: nonstructural protein NSI *status predicted <NSI>
F:134-1010/Product: nonstructural protein NSI *status predicted <NSI>
F:1316-1321/Region: nucleotide-binding motif A (P-loop)
F:1316-1321/Region: nucleotide-binding motif B
F:1320-1323/Region: DEXH motif
F:1867-2017/Product: nonstructural protein NSI *status predicted <NRA>
F:1867-2017/Product: nonstructural protein NSI *status predicted <NSI>
F:2018-3033/Product: nonstructural protein NSI *status predicted <NSI>
F:196,209,233,299,305.417,423,430,448,477,534,542,558,578,627,649,1091,1217,1259,203E
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 A.Molecule type: genomic RNA
A.Residues: 1-3033 < CORNA
A.Residues: 1-3034 < CORNA
A.Residues: 1-3034 < CORNA
A.Residues: 1-3034 < CORNA
A.Experimental source: 180.14te HC.-J6 from a Japanese individual
C.Superfamily: hepatitis C virus genome polyprotein
C.Superfamily: hepatitis C virus genome polyprotein
C.S. Freywords: ATP; glycoprotein; hydrolase; P-100p; polyprotein; serine proteinase; tra
F.-2-115/Product: capsid protein C *status predicted < CPC>
F.116-1191/Product: major envelope protein E *status predicted < CPC>
F.389-737/Product: major envelope protein E *status predicted < CPC>
F.380-733/Product: nonstructural protein NSI *status predicted < NSI>
F.310-11-1619/Product: nonstructural protein NSI *status predicted < NSI>
F.3116-1321/Region: nucleotide-binding motif B
F.330-1323/Region: DEXH motif
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   General polyprotein - hepatitis C virus (isolate HC-J6)
N;Contains: capsid protein C; envelope protein M; hepacivirin (EC 3.4.21.98) (nonstructural strotein NS4s; nonstructural protein NS4b; nonstructural protein NS5
C;Species: hepatitis C virus
C;Species: hepatitis C virus
C;Accession: JQ1303
R;Okamoto, H.; Okada, S.; Sugiyama, Y.; Kurai, K.; Iizuka, H.; Machida, A.; Miyakawa, J. Gen. Virol. 72, 2697-2704, 1991
A;Title: Nucleotide sequence of the genomic RNA of hepatitis C virus isolated from a A;Reference number: JQ1303
M;UD:92044440; PMID:1658196
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               F;1620-1866/Product: nonstructural protein NS4a #status predicted <N4A>
F;1867-2017/Product: nonstructural protein NS4b #status predicted <NAB>
F;2018-3033/Product: nonstructural protein NS5 #status predicted <N05>
F;2018-3037/Product: nonstructural protein NS5 #status predicted <N05>
F;196,209,234,305,325,417,423,430,448,477,534,542,558,578,627,649,1091,1217,1259,2038
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     79 IASPKGPVTQMYTNVDKDLVGWQAPQGSRSLTPCTCGSSDLYLVTRHADVIPVRRRGDSR 138
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        139 GSLLSPRPISYLKGSSGGPLICPAGHAVGIFRAAVSTRGVAKAVDFIPVESLETTMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             19 TAYAQOTRGEQGCQKTSHTGRDKNQVEGEVQIVSTATQTFLATSINGVLWTVYHGAGTRT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         DB 1; Length 3033;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Length 3033;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      30; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  29; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       66.0%; Score 673; DB 1;
68.7%; Pred. No. 4.6e-53;
tive 27; Mismatches 29,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               66.2%; Score 675; DB 1, 69.8%; Pred. No. 3e-53; iive 24; Mismatches
```

QY         53 TATQTELATSINGVLWTVYHGAGTRTIASPKGPVTQWYTNVDKDLVGWQAPQGSRSLIPC 112           Db         1006 TSTTRSMGTCVNGVMYTTYHGSNARTLAAQMGPVNSRWWSASDDVAVYPLPVGARCLEPC 1065           QY         113 TGGSSDLYLVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCP 161           Db         1066 KCQPGGVWVIRNDGALCHGTLGRTVELDLFAELCDFRGSSGSPLLCD 1112           QY         162 AGHAVGTRAVSTRGVRAVDFLPTURSELTTHRRSP 197           HIHI :	RESULT 14 H83144 H83144 probable aromatic acid decarboxylase PA4019 [imported] - Pseudomonas aeruginosa (strictorable aromatic acid decarboxylase PA4019 [imported] - Pseudomonas aeruginosa (strictorable i 5-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000 C;Accession: H83144 R;Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrener, P.; Hickey, M.J.; adman, S.; Yuan, Y.; Pham, X.Q.; Erwin, A.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, R.; I.; Lory, S.; Olson, M.V. Nature 406, 959-964, 2000 A;Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic pe A; Accession: H83144 A;Status: preliminary A;Molecule type: DNA A;Residues: lr209 <2709 <2709 <2709 <2709 <2709 <2709 <2709 <2709 <2709 <2709 <2709 <2709 <2709 <2709 <2709 <2709 <2709 <2709 <2709 <2709 <2709 <2709 <2709 <2709 <2709 <2709 <2709 <2709 <2709 <2709 <2709 <2709 <2709 <2709 <2709 <2709 <2709 <2709 <2709 <2709 <2709 <2709 <2709 <2709 <2709 <2709 <2709 <2709 <2709 <2709 <2709 <2709 <2709 <2709 <2709 <2709 <2709 <2709 <2709 <2709 <2709 <2709 <2709 <2709 <2709 <2709 <2709 <2709 <2709 <2709 <2709 <2709 <2709 <2709 <2709 <2709 <2709 <2709 <2709 <2709 <2709 <2709 <2709 <2709 <2709 <2709 <2709 <2709 <2709 <2709 <2709 <2709 <2709 <2709 <2709 <2709 <2709 <2709 <2709 <2709 <2709 <2709 <2709 <2709 <2709 <2709 <2709 <2709 <2709 <2709 <2709 <2709 <2709 <2709 <2709 <2709 <2709 <2709 <2709 <2709 <2709 <2709 <2709 <2709 <2709 <2709 <2709 <2709 <2709 <2709 <2709 <2709 <2709 <2709 <2709 <2709 <2709 <2709 <2709 <2709 <2709 <2709 <2709 <2709 <2709 <2709 <2709 <2709 <2709 <2709 <2709 <2709 <2709 <2709 <2709 <2709 <2709 <2709 <2709 <2709 <2709 <2709 <2709 <2709 <2709 <2709 <2709 <2709 <2709 <2709 <2709 <2709 <2709 <2709 <2709 <2709 <2709 <2709 <2709 <2709 <2709 <2709 <2709 <2709 <2709 <2709 <2709 <2709 <2709 <2709 <2709 <2709 <2709 <2709 <2709 <2709 <2709 <2709 <2709 <2709 <2709 <2709 <2709 <2709 <2709 <2709 <2709 <2709 <2709 <2709 <2709 <2709 <2709 <2709 <2709 <2709 <2709 <2709 <2709 <2709 <2709 <2709 <2709 <2709 <2709 <2	A; Experimental source: strain PAOI C; Genetics: PAOI9 C; Genetics: PAOI9 C; Genetics: A; Genetics: A; Gene. C; Genetics: A; Genetics: A; Gene. C; Genetics: A; Gene. C; Superfamily: dedF protein Query Match B: 4%; Score 85.5; DB 2; Length 209; Best Local Similarity 27.9%; Pred. No. 1.7; Matches 51; Conservative 16; Mismatches 61; Indels 55; Gaps 11;	OY 43 QVEGEVO-IVSTATQTFLATSINGVLWTVYHGAGTRTIASPKGFVTOWTT 91	QY       138 RGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVSTRGVAKAVDFIPVES 189         13	probable periplasmic serine proteinase DO (htrA-1) · syphilis spirochete C; Specias: Treponema pailidum subsp. pailidum (syphilis spirochete) C; Specias: Treponema pailidum subsp. pailidum (syphilis spirochete) C; Date: 24-Jul-1998 #sequence_revision 24-Jul-1998 #text_change 09-Dec-2002 C; Accession: B71284  R;Fraser, C.M.; Norris, S.J.; Weinstock, G.M.; White, O.; Sutton, G.G.; Dodson, R.; G rson, J.; Khalak, H.; Richardson, D.; Howell, J.K.; Chidambaram, M.; Utterback, T.; M they, L.; Weidman, J.; Smith, H.O.; Venter, J.C. Science 281, 375-388, 1998  A;Tille: Complete genome sequence of Treponema pallidum, the syphilis spirochete. A; Reference number: A71250; MUID:98332770; PMID:9655876  A;Reference number: A71250; MUID:98332770; PMID:965876  A;Status: preliminary; nucleic acid sequence not shown; translation not shown A;Residues: 1-398 <ccl></ccl>
Qy         19 TAYAOQTRGEDGCQKTSHTGFDKNOVEGEVQIVSTATQTFLATSINGVLWTVYHGAGTRT 78           1034 TAYAOQTRGLLGTIVVSMTGRDKTEQAGEIQVLSTVTQSFLGTTISGVLWTVYHGAGNKT 1093           Qy         79 IASPKGPVTQWTRVDKDEVGWQAPQGSRSLTPCTCGSSDLXIVTRHADVIPVRRRGDSR 138           1034 LAGSRGPVTQWTSSAEGDLVGWPSPPGTKSLEPCTCGAVDLXLVTRHADVIPVRRRGDSR 138           111111111111111111111111111111111111	RESULT 12 polyprotein - douroucouli hepatitis GB virus A polyprotein - douroucouli hepatitis GB virus A C; Species: douroucoul hepatitis GB virus A C; Species: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 17-Nov-2000 C; Accession: T08841 R; Erker, J.C.; Desai, S.M.; Leary, T.P.; Chalmers, M.L.; Montes, C.C.; Mushahwar, I.K. J. Gen. Virol. 79, 41-45, 1998 A, Title: Genomic analysis of two GB virus A variants isolated from captive monkeys. A, Reference number: 216486. MUID:98120818; PMID:9460920 A, Accession: T08841 A, Molecule type: mRNA A, Molecule type: mRNA A, Molecule type: mRNA A, Residues: 1-3065 < ERK> A, Cross-references: EMBL:AF023425; NID:92828599; PIDN:AAC40502.1; PID:92828600 C; Superfamily: hepatitis C virus genome Polyprotein C; Keywords: polyprotein	Query Match 24.6%; Score 251; DB 2; Length 3005; Best Local Similarity 34.1%; Pred: No. 2.6e-14; Matches 56; Conservative 29; Mismatches 69; Indels 10; Gaps 3; Qy 33 KTSHGGRONYGEROUVERPATCHEATSINGVLWATVYHGAGRETIASPKGPVTQMYTN 92	1055 PSDDVAVYPLPSGASCLEPCKCGTOSVWCIRNDGALCHGRLSKLVELDLPTEISDFRG 153 SSGGPLLCPAGHAVGIFRAAVSTRGVAKAVDFIPVES 189 1113 SSGSPILCDEGHVVGMA-VSVLHRGVKVTGVRYVKPWETLPKDS 1155	RESULT 13 108839 polyprotein - marmoset hepatitis GB virus A C;Species: marmoset hepatitis GB virus A C;Species: marmoset hepatitis GB virus A C;Date: 23-Jul-1999 #sequence_revision 23-Jul-1999 #text_change 17-Nov-2000 C;Accession: T08839 R;Erker, J.C.; Desai, S.M.; Leary, T.P.; Chalmers, M.L.; Montes, C.C.; Mushahwar, I.K. J. Gen. Virol. 79, 41-45, 1998 A;Title: Genomic analysis of two GB virus A variants isolated from captive monkeys. A;Reference number: 216486; MUID:98120818; PMID:9460920	A:Status: translated from GB/EMBL/DDBJ A:Rotaus: translated from GB/EMBL/DDBJ A:Molecule type: genomic RNA A:Rotaules: 1-2970 < CERK> A:Cross-references: EMBL:AF023424; NID:g2828597; PIDN:AAC40501.1; PID:g2828598 C;Superfamily: hepatitis C virus genome polyprotein C;Superfamily

```
A,Cross-references: GB:AE001248; GB:AE000520; NID:g3323074; PIDN:AAC65740.1; PID:g332308
A;Experimental source: strain Nichols
C;Genetian
A;Genetian
A;Genetian
C;Superfamily: Escherichia coli trypsin-like proteinase degS; GLGF domain homology; tryp
                                                                                                                             69; Gaps 10;
                                                                                                                                                                          91 INVDRD----LVGWQAPQGSRSLIPCICGSSDLYLVTRHADV-----178
                                                                                                                                                                                                                                       59 LATSINGVLW-----TVYHGAGTRTIASPKGPV------TQMY------90
                                                                                           Query Match

8.4%; Score 85.5; DB 2; Length 398;
Best Local Similarity 23.0%; Pred. No. 3.7;
Matches 44; Conservative 29; Mismatches 49; Indels 6
                                                                                                                                                                                                                                                                                                                                           180 KAVDF-IPVES 189
                                                                                                                                                                                                                                                                                                                                                                 | | :||::
255 SGVGFAVPVDT 265
                                                                                                                                                            ò
                                                                                                                                                                                  a 5 a
                                                                                                                                                                                                                                                                         O.Y
D.D
                                                                                                                                                                                                                                                                                                                                   οy
                                                                                                                                                                                                                                                                                                                                                                  qq
```

Search completed: August 30, 2003, 19:20:31 Job time: 17.2134 secs

```
GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.
```

OM protein - protein search, using sw model

(without alignments) 949.524 Million cell updates/sec August 30, 2003, 18:01:52; Search time 9.75674 Seconds Run on:

1 MKKKGSVVIVGRINLSGDTA......VAKAVDFIPVESLETIMRSP 197 US-09-965-594-20 1020 Title: Perfect score: Sednence:

127863 segs, 47026705 residues BLOSUM62 Gapop 10.0 , Gapext 0.5 Searched:

Scoring table:

Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0 Maximum DB seq length: 200000000

Post-processing: Minimum Match 0% Maximum Match 100% Listing first 45 summaries

SwissProt\_41:\* Database : Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Result		* Ouery			SUMMARIES	
<u>.</u>	Score	Match	Length	DB.	ID	Description
7	854.5	w.	3011	٦	POLG_HCV1	ч
~	848.5	83.2	3011	-	POLG_HCVH	P27958 h genome po
e	836.5	ď	3010	-	POLG_HCVTW	Þ
4	827.5	ij	3010	7	POLG_HCVJT	h genome
ď	823.5	ö	3010	~	POLG_HCVBK	h genome
9	823.5	ö	3010	-	POLG_HCVJA	h genome
7	675	ė.	3033	-	POLG_HCVJ8	h genome
œ	673	Ġ	3033	~	POLG_HCVJ6	h genome
σ,	87		321	٦	HHOA_ARATH	
10	85.5		209	Н	PAAD_PSEAE	
11	m	•	437	Н	DEG1_ARATH	
13			452	-	AAMP_HUMAN	
13	78.5		485	-	Y136_TREPA	
14	œ		764	-	ICCR_DROME	drosophile
15			1165	-	POL_GALV	
16	77.5	•	263	М	GRAK_MOUSE	
17	w	•	323	-	VPRT_SMRVH	
18	w	•	333	-	MOSA_RHIME	
19	92		401	~	FXH1_MOUSE	O88621 mus musculu
20	9.2		3411	-	POLG_YEFV1	P03314 y genome po
21	94		3411	-	POLG_YEFV2	
22		•	3414	-	POLG_TBEVW	
23	u,	•	248	-	TRY1_CHICK	
24	W 1		452	7	MLTD_ECOLI	_
25	75.5		2269	-	WDR9_HUMAN	
56			467	-	NX1B_BOVIN	
27	74.5		248	7	GRAD_MOUSE	P11033 mus musculu
28	4		3414	-	POLG_LANVT	P29837 1 genome po
53	74		911	-	TB11_NEIMB	neisseria
30	74	•	973	-	VP18_HUMAN	
31	7		3414	٦	POLG_TBEVH	99 t genom
32	73.5	7.2	248	ч	TRY2_CHICK	gallus ga
33	m	٠	264	Н	CTRL_HUMAN	3 homo sa

EMBL; M62321; AAA45676.1; -. PIR; A39166; GNWVC3.

PIR; A39166; GNWVC3. PDB; 1A1V; 16-FEB-99. PDB; 1HEI; 25-NOV-98.

P54856 ustilago ma		P27425 equus cabal	P59240 mus musculu	P28863 oryctolagus	Q28146 bos taurus	P07720 t genome po	Q04538 t genome po	P29326 hepatitis e	P33426 hepatitis e	
DPM1_USTMA	FLII_AQUAE	TRFE_HORSE	NPH4_MOUSE	MM03_RABIT	NX1A_BOVIN	POLG_TBEVS	POLG_POWVL	VST2_HEVBU	VST2_HEVPA	MPRI_BOVIN
4-	۱ ٦	-	-	~	-	-	-	~	-	н
301	443	206	1425	478	1530	3412	3415	099	099	2499
7.2	7.2	7.2	7.2	7.2	7.2	7.2	7.2	7.1	7.1	7.1
73.5	73.5	73.5	73.5	73	73	73	73	72.5	72.5	72.5

## ALIGNMENTS

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        PESCENCE FROM N.A.

ESCUENCE FROM N.A.

ESCUENCE FROM N.A.

MEDLINE-91172826; PubMed-1848704;

MEDLINE-91172826; PubMed-1848704;

Choo O.-L., Richman K.H., Han J.H., Berger K., Lee C., Dong C.,

Choo O.-L., Medina-Selby A., Barr P.J., Weiner A.J.,

Callegos C., Coit D., Medina-Selby A., Barr P.J., Weiner A.J.,

Callegos C., Coit D., Medina-Selby A., Barr P.J., Weiner A.J.,

Callegos C., Coit D., Medina-Selby A., Barr P.J., Weiner A.J.,

Callegos C., Coit D., Medina-Selby A., Barr P.J., Weiner A.J.,

Callegos C., Coit D., Medina-Selby A., Barr P.J., Weiner A.J.,

Callegos C., Coit D., Medina-Selby A., Barr P.J., Weiner A.J.,

Callegos C., Coit D., Medina-Selby A., Barr P.J.,

Callegos C., Coit D., Medina-Selby C., Coit D., Caller A., Caller A., Caller C., Coit C., Caller C., Coit C.
                                                                                                                                                                          01-AUG-1992 (Rel. 23, Created)
01-AUG-1992 (Rel. 23, Last sequence update)
15-SEP-2003 (Rel. 24, Last annotation update)
66-1992 (Rel. 25, Last annotation update)
66-1992 (Rel. 25, Last annotation update)
67-1092 (Gontains: Capala protein C (Core protein) (P22);
67-1093 (GPP3) (GPP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    -1- SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A LIPOPMOTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS: PROTEIN AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MRNA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        -! - SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY S29.
                                                                                                       PRT; 3011 AA
                                                                                                       STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Hepacivirus.
NCBI_TaxID=11104;
                                                                                                       POLG_HCV1
RESULT 1
POLG_HCV1
```

```
PRT; 3011 AA
                                                                                                                                          174 STRGVAKAVDFIPVESLETTMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          of unwinding.";
Structure 6:89-100(1998)
                                                                                                                                                                                                              STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ACTIVATION OF NS3.
                                                                                                                                                                                                                                                                                                                                                                                 NCBI_TaxID=11108;
                                                                                                                                                                                                                                                                                                                                                                                                                                           Prince A.M.;
                                                                                                                                                                                                                                                                                                                                                                        Repacivirus
                                                                                                                                                                                                           POLG_HCVH
P27958;
                                                                                                                                                                                       RESULT 2
POLG_HCVH
                                                                                                       g
                                                                                                                                                    a
                                  ò
                                                         a
                                                                                ò
                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                              CELLULAR ANINOPEPTIDASE.
CAPSID PROTEIN ( POTENTIAL).
MATAIX PROTEIN ( POTENTIAL).
MAJOR ENVELOPE PROTEIN E (POTENTIAL).
NONSTRUCTURAL PROTEIN NS1/E2 (POTENTIAL).
PROTEASE/HELICASE NS3 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4A (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4A (POTENTIAL).
NANA DIRECTED RNA POLYMERASE (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   9; Gaps
                                                                                                                                                                                                                                                                                        Polyprotein; Glycoprotein; Transferase; RNA-directed RNA polymerase; Core protein; Coat protein; Envelope protein; Helicase; ATP-binding; Transmembrane; Nonstructural protein; Hydrolase; Serine protease;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SIMILARITY).
SIMILARITY).
SIMILARITY).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  (POTENTIAL). (POTENTIAL). (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                (POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           (POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             DB 1; Length 3011;
                                                                                                                                                                                                                                                                                                                                     REMOVED FROM CAPSID PROTEIN C
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  9; Mismatches 17; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      65F8C9447FCE5AF9 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CHARGE RELAY SYSTEM (BY
CHARGE RELAY SYSTEM (BY
CHARGE RELAY SYSTEM (BY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (GLCNAC. . . .
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            (GLCNAC. .
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        (GLCNAC. .
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        (GLCNAC. .
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Score 854.5; DB 1
Pred. No. 2.5e-72;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           (GICNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         N-LINKED (GLCNAC.
N-LINKED (GLCNAC.
N-LINKED (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            N-LINKED (GLCNAC
N-LINKED (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ATP (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                 POTENTIAL.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   N-LINKED
                                          Interpro; IRR000745; HCV_NS3.
Interpro; IPR000745; HCV_NS4.
Interpro; IPR000745; HCV_NS4.
Interpro; IPR000749; HCV_NS4.
Interpro; IPR0001669; HCV_RS4.
Interpro; IPR0001650; Helicasc_C.
Interpro; IPR0001650; Helicasc_C.
Interpro; IPR000095; RNA_POL_DS_PS.
Interpro; IPR000095; RNA_POL_DS_PS.
Interpro; IPR000094; HCV_NS3; I.
Pfam; PF01509; HCV_NS3; I.
Pfam; PF01001; HCV_NS4; I.
Pfam; PF01001; HCV_NS4; I.
Pfam; PF01001; HCV_NS4; I.
Pfam; PF00001; HOV_NS5; I.
Pfam; PF00001; HOV_NS5; I.
Pfam; PF00001; HOV_NS5; I.
Pfam; PF00001; HOV_NS5; I.
Pfam; PF000001; HOV_NS5; I.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           BOX
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          DECH
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               645 645 N-LIN
2041 2041 N-LIN
2240 2240 N-LIN
2364 2364 N-LIN
2362 2769 N-LIN
3011 AA: 327197 MW; G
HCV_core.
HCV_env.
HCV_NS1.
HCV_NS2.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          tch 83.8%; al Similarity 82.8%; 169; Conservative
                                                                                                                                                                                                                                                                               SMART; SM00487; DEXDC;
                                                                                                                                                                                                                                                                                                                                                                                                                                                             369
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                165
                                  IPR002518;
                                                                                                                                                                                                                                                                                                                            3D-structure.
INIT_MET
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Query Match
Best Local S
Matches 169
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CARBOHYD
CARBOHYD
            InterPro;
                         InterPro;
                                  InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ACT_SITE
ACT_SITE
ACT_SITE
NP_BIND
SITE
                                                                                                                                                                                                                                                                                                                                                                                                                                                     CHAIN
TRANSMEM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CARBOHYD
                                                                                                                                                                                                                                                                                                                                                          CHAIN
CHAIN
CHAIN
                                                                                                                                                                                                                                                                                                                                                                                             CHAIN
CHAIN
CHAIN
CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                           CHAIN
```

3 KKGSVVIVGRIN-----LSGDTAYAQQTRGEQGCQKTSHTGRDKNQVEGEVQIVST 53

ò

```
113
                                                                          .065 AAQTFLATCINGVCWTVYHGAGTRTIASPKGPVIQMYTNVDQDLVGWPAPQGSRSLTPCT 1124
                                                                                                                                                     114 CGSSDLYLVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAV 173
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    01-A0G-1992 (Rel. 23, Created)
01-A0G-1992 (Rel. 23, Last sequence update)
01-A0G-1992 (Rel. 23, Last sequence update)
01-A0G-1992 (Rel. 24, Last annotation update)
Genome polyprotein (Contains: Capsid protein C (Core protein) (P22);
Envelope glycoprotein El (GP32) (GP35); Envelope glycoprotein E2
(GP68) (GP70) (NS1); Protein P7; Nonstructural protein NS2 (P21)
(EC 3.4.21.98); Nonstructural protein NS48 (P4); Nonstructural protein NS48 (P27); Nonstructural protein NS58 (P56); Nonstructural protein NS58 (P65); N
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    X-RAY CRYSTALLOGRAPHY (2.1 ANGSTROMS) OF 1207-1657.
MEDLINE-97331322; PubMed-9187654;
Yao N., Hesson T., Cable M., Hong Z., Kwong A.D., Le H.V., Weber P.C.;
"Structure of the hepatitis C virus RNA helicase domain.";
Nat. Struct. Biol. 4:463-467(1997).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          -: FUNCTION: PROTESSE NS2 IS RESPONSIBLE FOR THE CLEAVAGE OF NS2-NS3-I- FUNCTION: PROTESSE NS3 IS RESPONSIBLE FOR THE CLEAVAGE OF NS3-NS4A, NS4A-NS4B, NS4B-NS5A, AND NS5A-NS5B, NS4A-NS4B, NS4A-NS4B, NS4B-NS5B, AND NS5A-NS5B, PUNCTION: NS4A FORMS A COMPLEX WITH NS3 AND IS ESSENTIAL FOR THE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               -: FUNCTION: NSSA SEEMS TO HAVE A TRANSCRIPTIONAL ACTIVATORY ROLE.
-: FUNCTION: NSSB IS A RNA-DEPENDENT RNA POLYMERASE THAT PLAYS AN
ESSEMTIAL ROLE IN THE VIROS REPLICATION.
-: CATALYTICA ACTIVITY: Hydrolysis of four peptide bonds in the viral
precursor polyprotein, commonly with Asp or Glu in the P6
position, Cys or Thr in P1 and Ser or Ala in P1'.
-: CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
54 ATQTFLATSINGVLWTVYHGAGTRTIASPKGPVTQMYTNVDKDLVGWQAPQGSRSLTPCT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          {RNA}(N).
SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A
LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS: E1
AND E2. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MRNA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS) OF 1192-1657.
MEDLINE-98154321; PubMed-9493270;
Kim J.L., Morgenstern K.A., Griffith J.P., Dwyer M.D., Thomson J.A.,
Murcko M.A., Lin C., Caron P.R.;
"Hepatitis C virus NS3 RNA helicase domain with a bound
oligonucleotide: the crystal structure provides insights into the mon
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ပ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       "Genomic structure of the human prototype strain H of hepatitis virus: comparison with American and Japanese isolates."; Proc. Natl. Acad. Sci. U.S.A. 88:10292-10296(1991).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SEQUENCE FROM N.A.
MEDLINE-92052256; PubMed-1658800;
Inchauspe G., Zebedee S., Lee D.H.H., Sugitani M., Nasoff M.,
```

က

```
(POTENTIAL)
(POTENTIAL)
(POTENTIAL)
(POTENTIAL)
(POTENTIAL)
(POTENTIAL)
(POTENTIAL)
(POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Length 3011;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           327142 MW; 772CBB29CCD94753 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Score 848.5; DB 1;
Pred. No. 9.2e-72;
                             (GLCNAC.
(GLCNAC.
(GLCNAC.
(GLCNAC.
(GLCNAC.
(GLCNAC.
                                                                                                                                                                                                                            (GLCNAC.
(GLCNAC.
(GLCNAC.
                             N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   83.2%;
81.9%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1417
1420
1420
1420
1443
1463
1463
1468
1469
1501
1501
1511
1511
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1377
1380
1385
1393
1409
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            353
361
366
368
375
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     11312
113212
1133213
1133213
1133213
1133213
1133213
1133213
113413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
11413
1141
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1497
1507
1511
1511
1532
1532
1555
1555
1606
1622
1622
1637
1635
1635
1635
1635
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           TURN

STRAND

TURN

HELIX

TURN

TURN

TURN

STRAND

STRAND

STRAND

STRAND

STRAND

STRAND

STRAND

TURN

STRAND

STRAND

STRAND

STRAND

STRAND

STRAND

TURN

TURN

STRAND

STRAND

STRAND

STRAND

HELIX

TURN

HELIX

HELIX

HELIX

HELIX

HELIX

HELIX

HELIX

TURN

HELIX

STRAND

STRAND

STRAND

STRAND

HELIX

TURN

HELIX

TURN

HELIX

TURN

HELIX

HELIX

TURN

HELIX

HELIX

TURN

HELIX

HELIX

HELIX

HELIX

TURN

HELIX

HELIX
                             CARBOHYD
TURN
HELIX
TURN
HELIX
TURN
HELIX
TURN
STRAND
TURN
STRAND
TURN
TURN
STRAND
TURN
STRAND
TURN
HELIX
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Ouery Match
Best Local
     CARBOHYD
     This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation. The European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                 EMBL, M67463, AA44534.1; --

BY PIR, A36844; GNWYCH.

BY PDB, 1HEI: 25-NOY-98.

BY PDB, 1HEI: 25-NOY-98.

BY RENOPS, 529-001; --

BY INTERPORT OF TOURS; --

BY INTERPORT INTOURNED.

BY INTERPORT I
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CHARGE RELAY SYSTEM (BY SIMILARITY).
CHARGE RELAY SYSTEM (BY SIMILARITY).
CHARGE RELAY SYSTEM (BY SIMILARITY).
ATP (POTENTIAL).
DECH BOX.
N-LINKED (GLCNAC. ..) (POTENTIAL).
N-LINKED (GLCNAC. ..) (POTENTIAL).
N-LINKED (GLCNAC. ..) (POTENTIAL).
PTM: THE STRUCTURAL PROTEINS C, E1 AND E2 ARE PRODUCED BY PROFELYTIC PROCESSING BY THE HOST SIGNAL PEPTIDASES.
SIMILARITY: THE NS2 PROTEASE BELONGS TO PEPTIDASE FAMILY U39. SIMILARITY: THE NS3 PROTEASE BELONGS TO PEPTIDASE FAMILY S29.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     REMOVED FROM CAPSID PROTEIN C
CELLULAR AMINOPEPTIDASE.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    NONSTRUCTURAL PROTEIN NS2.
PROTEASE/HELICASE NS3.
NONSTRUCTURAL PROTEIN NS4A.
NONSTRUCTURAL PROTEIN NS4B.
NONSTRUCTURAL PROTEIN NS5B.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CAPSID PROTEIN C.
ENVELOPE GLYCOPROTEIN E1.
ENVELOPE GLYCOPROTEIN E2.
PROTEIN P7.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          191
383
746
809
1026
1025
11657
11657
1107
1107
1107
11107
11107
11107
11107
11107
11107
11107
11107
11107
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CCHAIN
CCHAIN
CCHAIN
CCHAIN
CCHAIN
CCHAIN
CCHAIN
CHAIN
```

us-09-965-594-20.rsp

```
ACT_SITE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           INITUMET
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          IRANSMEM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CHAIN
CHAIN
CHAIN
CHAIN
CHAIN
CHAIN
  SOLUTION OF THE PROPERTY OF TH
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation he Burnepan Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as its content is no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (Sec http://www.isb-sib.ch/announce/
                                                                                                                                         1125 GGSSDLYLVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPIGHAVGLFRAAV 1184
                                                           54 ATQTFLATSINGVLWTVYHGAGTRTIASPKGPVTQMYTNVDKDLVGWQAPQGSRSLTPCT 113
                                                                                                                                                                                                114 CGSSDLYLVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAV 173
                                                                                                                                                                                                                                                                                                                                                                                                                                                  01-APR-1993 (Rel. 25, Created)
01-APR-1993 (Rel. 25, Last sequence update)
01-APR-1993 (Rel. 25, Last sequence update)
01-APR-1993 (Rel. 42, Last annotation update)
Genome polyprotein (Contains Capsid protein C (Core protein) (P22);
Envelope glycoprotein El (GP32) (GP35); Envelope glycoprotein E2
(GP68) (GP70) (N81); Protein P1 Nonstructural protein NS2 (P21)
(EC 3.4.21.98); Nonstructural protein NS48 (P4); Nonstructural protein
NS48 (P27); Nonstructural protein NS5A (P56); Nonstructural protein
NS5B (P66) (P70) (RNA-directed RNA polymerase) (EC 2.7.7.48)].
Hepatitis C virus (1solate Taiwan) (HCV).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           3 KKGSVVIVGRIN-----LSGDTAYAQOTRGEQGCQKTSHTGRDKNQVEGEVQIVST
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (RNA)(N).
SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A LIPOPROTEIN ENVELOPE CONSISTS OF TWO PROTEINS.
PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MRNA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY $29.
  Indels
  18;
10; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                   174 STRGVAKAVDFIPVESLETTMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          EMBL; M84754; -; NOT_ANNOTATED_CDS.
PIR; A40244; GNWVTW.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  MEROPS; 229.001; -.
MEROPS; 039.001; -.
InterPro; IPR001410; DEAD.
Matches 167; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                   STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1N64; 25-FEB-03.
1NS3; 08-APR-98.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Hepacivirus.
NCBL_TaxID=31645;
                                                                                                                                                                                                                                                                                                                                                                                                                 POLG_HCVTW
P29846;
                                                                                                                                                                                                                                                                                                                                                                                                                   a
                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                            à
```

```
ï
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             REMOVED FROM CAPSID PROTEIN C BY THE CELLULAR AMINOEPTIDASE.
CORE PROTEIN (POTENTIAL).
MATRIX PROTEIN (POTENTIAL).
MAJOR ENVELOPE PROTEIN E (POTENTIAL).
NONSTRUCTURAL PROTEIN NSI/EZ (POTENTIAL).
PROYEASE/HELICASE NS3 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4A (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4A (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4B (POTENTIAL).
RNA-DIRECTED RNA POLYMERASE (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Polyprotein; Glycoprotein; Transferase; RNA-directed RNA polymerase; Core protein; Coat protein; Envelope protein; Helicase; ATP-binding; Transmembrane; Nonstructural protein; Hydrolase; Serine protease;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SIMILARITY)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ä
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          82.0%; Score 836.5; DB 1; Length 3010;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         2788 N-LINKED (GLCNAC. . .) (POTI
AA; 327047 NW; AAD267D55CDFE215 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CHARGE RELAY SYSTEM (BY
CHARGE RELAY SYSTEM (BY
CHARGE RELAY SYSTEM (BY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           N-LINKED (GLCNAC. ..)
N-LINKED (GLCNAC. ..)
N-LINKED (GLCNAC. ..)
N-LINKED (GLCNAC. ..)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      N-LINKED (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ATP (POTENTIAL).
DECH BOX.
N-LINKED (GLCNAC.
N-LINKED (GLCNAC.
N-LINKED (GLCNAC.
N-LINKED (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Pred. No. 1.3e-70;
18; Mismatches 17
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   N-LINKED (GLCNAC.
                                                                                                                                      "" NS4a.
"" NS4a.
"" NS4b.
"" NS5a.
"" InterPro; IPR002166; HCV_NS5a.
"" InterPro; IPR007095; RNA_POL_DS_PS.
"" Pfam; PP01543; HCV_capsid; I. Pfam; PP01543; HCV_capsid; I. Pfam; PP01542; HCV_capsid; I. Pfam; PP01542; HCV_capsid; I. Pfam; PP01550; HCV_capsid; I. Pfam; PP01560; HCV_capsid; I. Pfam; PP01560; HCV_capsid; I. Pfam; PP01560; HCV_capsid; I. Pfam; PP01560; HCV_capsid; I. Pfam; PP01550; HCV_capsid; III NFAM; PP01550; HCV_capsid;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                Pfam, PF01560; HCV_NS1; 1.
Pfam; PF01550; HCV_NS1; 1.
Pfam; PF01006; HCV_NS2; 1.
Pfam; PF01006; HCV_NS4a; 1.
Pfam; PF01001; HCV_NS4b; 1.
Pfam; PF01001; HCV_NS5b; 1.
Pfam; PF01050; HCV_NS5a; 1.
Pfam; PF00999; Viral_RGRP; 1.
ProDom; PD186062; HCV_NS1; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            78.48;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      2788
                                                                                                                            IPRO04109; H
IPR000745; H
IPR001490; H
IPR002868; H
IPR002521;
IPR002519;
                                                                 IPR002531;
IPR002518;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Best Local Sim
Matches 160;
                                  InterPro;
InterPro;
InterPro;
                                                                                                                                   InterPro;
InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ACT_SITE
ACT_SITE
NP_BIND
SITE
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CARBOHYD
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CARBOHYD
```

```
REMOVED FROM CAPSID PROTEIN C BY THE CELLULAR AMINOEPTIDASE.
CAPSID PROTEIN C (POTENTIAL).
MAJOR ENVELOPE PROTEIN E (POTENTIAL).
MAJOR ENVELOPE PROTEIN NSI/E2 (POTENTIAL).
NONSTRUCTURAL PROTEIN NSI/E2 (POTENTIAL).
PROTEASE/HELICASE NS3 (POTENTIAL).
NONNSTRUCTURAL PROTEIN NS4A (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4A (POTENTIAL).
RNA-DIRECTED RNA POLYMERASE (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               KKGSVVIVGRIN-----LSGDTAYAQQTRGEQGCQKTSHTGRDKNQVEGEVQIVST
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Polyprotein; Glycoprotein; Transferase; RNA-directed RNA polymerase; Core protein; Coat protein; Envelope protein; Helicase; ATP-binding; Transmembrane; Nonstructural protein; Hydrolase; Serine protease;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CHARGE RELAY SYSTEM (BY SIMILARITY)
CHARGE RELAY SYSTEM (BY SIMILARITY)
CHARGE RELAY SYSTEM (BY SIMILARITY)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (POTENTIAL) (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     DB 1; Length 3010;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              POTENT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           18; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            94A1C77435D642BB
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Score 827.5; DB 1
bred. No. 8.9e-70;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                N-LINKED (GLCNAC.
N-LINKED (GLCNAC.
N-LINKED (GLCNAC.
N-LINKED (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (GICNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           (GICNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   N-LINKED (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ATP (POTENTIAL)
DECH BOX.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         20; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    POTENTIAL.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       N-LINKED
N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              N-LINKED
                                                                                                                                                   1 IPR002868; HCV_NS5a.; IPR002166; HCV_RGRP.; IPR007095; RNA_DOL_DS_PS.; IPR007094; RNA_DOL_PSVIR.; HCV_Capsid; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              326573 MW;
                                                                                                 HCV_NS3.
HCV_NS4a.
HCV_NS4b.
                                                            HCV_NS1.
                                                                                                                                                                                                                                                                                                                                                                                        Pfam; PF01506; HCV_NS5a; 1.
Pfam; PF00271; helicase_C; 1.
Pfam; PF00998; Viral_RdRP; 1.
ProDom; PD186062; HCV_NS1; 1.
                                          HCV_env.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   81.1%;
77.0%;
                                                                                                                                                                                                                                                                                                                               HCV_NS3; 1
HCV_NS4a;
HCV_NS4b;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Ouery Match 81.1
Best Local Similarity 77.0
Matches 157; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SMART; SM00487; DEXDC;
                                                                                                                   IPR000745; I
      IPR002522;
                        IPR002521;
                                                                              IPR002518;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              3010 AA;
                                                                                                                                                                                                                                                                                                                                                   Pfam; PF01006; H
Pfam; PF01001; H
Pfam; PF01506; H
                                                                                                                                                                                                                                                                                           PF01560;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    D-structure.
                                                                                                 InterPro;
InterPro;
InterPro;
                                                                                                                                                           interPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ACT_SITE
NP_BIND
SITE
                                                                              InterPro;
                                                                                                                                                                              InterPro;
                                                                                                                                                                                              nterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CARBOHYD
CARBOHYD
                        InterPro;
                                          InterPro;
                                                              InterPro;
                                                                                                                                                                                                                    nterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ACT_SITE
ACT_SITE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  INIT MET
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              LRANSMEM
                                                                                                                                                                                                                                      Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CHAIN
CHAIN
CHAIN
CHAIN
CHAIN
CHAIN
                                                                                                                                                                                                                                                                                             Pfam;
                                                                                                                                                                                                                                                                                                                                   Pfam;
                                                                                                                                                                                                                                                          Pfam;
    à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation-the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                        1125 CGSSDLYLVTRHADVIPVRRKGDSRGSLLSPRPISYLKGSSGGPLLCPSGHVVGIFRAAV 1184
                                                                                                                                                       114 CGSSDLYLVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAV 173
                                                                                                                                                                                                                                                                                                                                                                                                    01-APR-1993 (Rel. 25, Created)
01-APR-1993 (Rel. 25, Last sequence update)
15-SEP-2003 (Rel. 25, Last sequence update)
6600me polyprotein [Contains: Capsid protein C (Core protein) (P22);
6700me polyprotein El (GP32) (GP35); Envelope glycoprotein E2
(GP68) (GP00) (NSI); Protein P7; Nonstructural protein NS2 (P21)
(EC 3.4.21.98); Nonstructural protein NS48 (P4); Nonstructural protein NS58 (P27); Nonstructural protein NS58 (P56); (P70) (RNA-directed RNA P0)/merase) (EC 2.7.7.48)].
                                                                                                   SEQUENCE FROM N.A.
MEDILINE-2925714; DUDMed-1318627;
MEDILINE-2925714; DEDMED-1318627;
Tanaka T., Kato N., Nakagawa M., Ootsuyama Y., Cho M.J.,
Nakazawa T., Hijikata M., Ishimura Y., Shimotchno K.;
"Molecular cloning of hepatitis C virus genome from a single Japanese
carrier: sequence variation within the same individual and among
infected individuals.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               VITUE RES. 23:99-53(1992).

-!- FUNCTION: THE SMALL PROTEINS NSZA, NSZB, NSAA AND NS4B ARE HYDROPHOBIC, SUGGESTING A POSSIBLE MEMBRANE-RELATED FUNCTION.

NS3 AND NS5 MAY PLAY A ROLE IN THE VIRAL RNA REPLICATION.

-!- CATALYTIC ACTIVITY: Hydrolysis of four peptide bonds in the viral precursor polyprotein, commonly with Asp or Glu in the P6 postition, Cys or Thr in Pl and Ser or Ala in Pl'.

-!- CATALYTIC ACTIVITY: N nucleoside triphosphate - N diphosphate +
3 KKGSVVIVGRIN-----LSGDTAYAQQTRGEQGCQKTSHTGRDKNQVEGEVQIVST
                                                                              ATQTFLATS INGVLWTVYHGAGTRT IASPRGPVTQMYTNVDKDLVGWQAPQGSRSLTPCT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (RNA)(N).
SUBDINI: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY
LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
PROTEIN M AND EXCOPPOTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF
PROTEIN C. AND MRNA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                    174 STRGVAKAVDFIPVESLETTMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        EMBL; D11168; BAA01943.1; -. PIR; A45573; A45573.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           InterPro; IPR001410; DEAD.
                                                                                                                                                                                                                                                                                                                                                                      STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1A10; 25-MAR-98.
1JXP; 14-JAN-98.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     MEROPS; $29.001; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             NCBI_TaxID=31642;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Hepacivirus
                                                                                                                                                                                                                                                                                                                                                                    POLG_HCVJT
                                                                                                                                                                                                                                                                                                                                                   POLG_RCVJT
                                                                                                                                                                                                                                                                                                                                                                        A PART A 
                                                                                                                                                                                        g
                                                                                                                                                                                                                                                              q
                                                                                                                                                                                                                                ò
                                                                                                                                                     οý
                                        合
                                                                              ò
                                                                                                         8
```

23

9; Gaps

us-09-965-594-20.rsp

```
1005 RRGREILLGPADSIEGGGWRLLAPITAYAQOTRGLIGCIVTSLTGRDKNQVEGEVQVVST 1064
                                                     X. REDILINE-9227846; PUDMed-9568991;
X. MEDLINE-99227846; PUDMed-9568991;
Yan Y., Li Y., Munshi S., Sardana W.,
Stelnkuehler C., Tomei L., de Francesco R., Kuo L.C., Chen Z.;
T. Complex of NS3 protease and NS4A peptide of BK strain hepatitis C
YIUS: a 2.2-A resolution structure in a hexagonal crystal form,";
T. POTALIA SOL, 7:837-847(1998)
C. 1- FUNCTION: THE SMALL PROTEINS NS2A, NS2B, NS4A AND NS4B ARE
HYDROPHOBIC. SUGGESTING A POSSIBLE MEMBRANE-RELATED FUNCTION.
NS3 AND NS5 MAY PLAY A ROLE IN THE VIRAL RNA REPLICATION.
C. 1- CATALYTIC ACTIVITY: Hydrolysis of four peptide bonds in the viral prosition, Cys or Thr in Pl and Ser or Ala in Pl.
C. -1- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
                                     ATQTFLATSINGVLWTVYHGAGTRIIASPKGPVTQMYTNVDKDLVGWQAPQGSRSLTPCT 113
                                                                                       114 CGSSDLYLVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAV 173
                                                                                                                                                                                                                                                             01-005 (Rel. 23, Created)
01-005 (Rel. 23, Last sequence update)
15-SEP-2003 (Rel. 42, Last nonotation update)
15-SEP-2003 (Rel. 42, Last annotation update)
16-Ontains: Capsid protein C (Core protein) (P22);
Envelope glycoprotein El (GP32) (GP35); Envelope glycoprotein E2
(GP68) (GP70) (N81); Protein P7; Nonstructural protein NS2 (P21)
(EC 3.4.22.-); Protease/helicase NS3 (P70) (Hepaclvirin)
(EC 3.4.21.98); Nonstructural protein NS5A (P4); Nonstructural protein
NS4B (P27); Nonstructural protein NS5A (P65); Monstructural protein
NS5B (P66) (P70) (Rna-directed RNA polymerase) (EC 2.7.7.48)].
Hepatitis C virus (isolate BNS) (HCV)
Viruses; SSRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 X-RAY CRYSTALLOGRAPHY (2.4 ANGSTROMS) OF 1027-1215.
MEDILINE-97015088: PubMed-8861916;
LOVE R.A., Parge H.E., Wickersham J.A., Hostomsky Z., Habuka N.,
MOOMAW E.W., Adachi T., Hostomska Z.;
The crystal structure of hepatitis C virus NS3 proteinase reveals a trypsin-like fold and a structural zinc binding site.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Takamizawa A., Mori C., Fuke I., Manabe S., Murakami S., Fujita J., Onishi E., Andoh T., Yoshida I., Okayama H.; *** Andoh T., Yoshida I., Okayama H.; **** Structure and organization of the hepatitis C virus genome isolated
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  MEDLINE-9623524; PubMed-8647104;
Borcowski P., Heiland M., Oehlmann K., Becker B., Korneteky L.;
"Non-structural protein 3 of hepatitis C virus inhibits
phosphorylation mediated by cAMP-dependent protein kinase.";
Eur. J. Biochem. 237:611-618(1996).
                                                                                                                                                                                                                                        Ş
                                                                                                                                          SEQUENCE FROM N.A.
MEDLINE-91140698; PubMed-1847440;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               J. Virol. 65:1105-1113(1991).
                                                                                                                                                                                                                                      STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SEQUENCE OF 1487-1500.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    from human carriers.
                                                                                                                                                                                                                                                                                                                                                                                                                                        NCBI_TaxID-11105;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (RNA) (N).
                                                                                                                                                                                                                                      POLG_HCVBK
P26663;
                                       24
                                                                                                                                                                                                                                        g
                                                                                                                                                                    g
                                                                                      ò
                                                                                                                                              ò
                                                              a
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bloinformatics Institute. There are no restrictions on its mes by non-profil institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CELLULAR AMINOPEPIDAGE.

CAPSID PROTEIN C (POTENTIAL).

MAJOR ENVELOR (POTENTIAL).

NONSTRUCTURAL PROTEIN E (POTENTIAL).

NONSTRUCTURAL PROTEIN NSI/PE2 (POTENTIAL).

PROTEASE/HELICASE NS3 (POTENTIAL).

NONSTRUCTURAL PROTEIN NS4A (POTENTIAL).

NONSTRUCTURAL PROTEIN NS4A (POTENTIAL).

NONSTRUCTURAL PROTEIN NS4B (POTENTIAL).

RNA-DIRECTED RNA POLYMERASE (POTENTIAL).
SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Polyprotein; Glycoprotein; Transferase; RNA-directed RNA polymerase; Core protein; Coat protein; Envelope protein; Helicase; ATP-binding; Transmembrane; Nonstructural protein; Hydrolase; Serine protease;
         LIPOPROTEIN ENVELOPE, THE ENVELOPE CONSISTS OF TWO PROTEINS:
PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF
PROTEIN C AND MRNA.
                                                                         SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY S29.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               REMOVED FROM CAPSID PROTEIN C
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CHARGE RELAY SYSTEM.
CHARGE RELAY SYSTEM.
CHARGE RELAY SYSTEM.
ATP (POTENTIAL).
DECH BOX.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     InterPro; IPR007095; RNA_pol_DS_PS.
InterPro; IPR007094; RNA_pol_PSvir.
Pfam: PF01543; HCV_capsid; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      HCV_capsid.
HCV_core.
HCV_env.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Pfam; PF01542; HCV_core; 1.
Pfam; PF01542; HCV_core; 1.
Pfam; PF01560; HCV_MS1; 1.
Pfam; PF01538; HCV_MS2; 1.
Pfam; PF01504; HCV_MS3; 1.
Pfam; PF01006; HCV_MS4a; 1.
Pfam; PF01001; HCV_MS4b; 1.
Pfam; PF01001; HCV_MS4b; 1.
Pfam; PF01001; HCV_MS5a; 1.
Pfam; PF01001; HCV_MS5a; 1.
Pfam; PF01001; HCV_MS5a; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             HCV_NS1.
HCV_NS2.
HCV_NS3.
HCV_NS4a.
HCV_NS4b.
HCV_NS5a.
                                                                                                                                                                                                                                                          EMBL; M58335; AAA72945.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       383
729
729
1006
1615
1862
2013
3010
369
1107
11107
11165
11137
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SMART; SM00487; DEXDC;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              InterPro; IPR002518; HG InterPro; IPR004109; HG InterPro; IPR000745; HG InterPro; IPR001490; HG
                                                                                                                                                                                                                                                                                           1A1Q: 25-MAR 98.
1JXP: 14-JAN-98.
1US3: 08-APR-98.
1C2P: 15-NOV-00.
1CSJ: 08-NOV-99.
1GX5: 09-APR-02.
1GX6: 10-APR-02.
1QUV: 26-JUN-00.
80HM: 20-APR-99.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           IPR002519;
IPR002531;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       InterPro; IPR001410;
InterPro; IPR002522;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            IPR002521;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            IPR002868;
                                                                                                                                                                                                                                                                               GNWVTC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       U39.00
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              3D-structure.
INIT_MET
                                                                                                                                                                                                                                                                             A38465;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       TRANSMEM
ACT_SITE
ACT_SITE
ACT_SITE
ACT_SITE
NP_BIND
SITE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  MEROPS;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 MEROPS;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CHAIN
CHAIN
CHAIN
CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CHAIN
                                                                                                                                                                                                                                                                                                                                  PDB;
PDB;
PDB;
PDB;
```

```
114 CGSSDLYLVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAV 173
                                                                                                                                                                                                                                                                                                                                                                                  53
                                                                                                                                                                                                                                                                                                                                                                                  3 KKGSVVIVGRIN-----LSGDTAYAQQTRGEQGCQKTSHTGRDKNQVEGEVQIVST
                                                                                                                                                                                                                                                                                                                                                                                                            54 ATQTFLATSINGVLWTVYHGAGTRTIASPKGPVTQMYTNVDKDLVGWQAPQGSRSLTPCT
                                                                                                                                                                                                                                                                                                                                                                     Gaps
                                                         (POTENTIAL).
(POTENTIAL).
(POTENTIAL).
(POTENTIAL).
(POTENTIAL).
                                (POTENTIAL).
(POTENTIAL).
(POTENTIAL).
                                                                                         (POTENTIAL). (POTENTIAL). (POTENTIAL).
                                                                                                            (POTENTIAL).
(POTENTIAL).
(POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                     6
                                                    (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                       Length 3010;
                                                                                                                                                                                                                                                                                                                                                                     Indels
                                                                                                                                                                                                                                                                                                                                           327189 MW; F8422D5ECCFDFD9C CRC64;
                                                                                                                                                                                                                                                                                                                                                       80.7%; Score 823.5; DB 1;
76.5%; Pred. No. 2.1e-69;
Live 21; Mismatches 18;
                                (GLCNAC.
(GLCNAC.
(GLCNAC.
                                                                            (GLCNAC.
(GLCNAC.
(GLCNAC.
                                                                (GLCNAC.
                                                                                               (GLCNAC.
(GLCNAC.
(GLCNAC.
                                                                                                                  (GLCNAC.
                                                    (GLCNAC.
                    (GLCNAC
                           GECNAC
N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                       174 STRGVAKAVDFIPVESLETTMRSP 197
                                                                                                                                                                                                                                                                                                                                                                     Conservative
                                                                                                                                                                                                                                                                                                     1186
1188
1197
1202
1204
                                                                                                                                                  1063
1074
1076
1081
1087
1092
1094
1103
                                                                                                                                                                                                                                                        11144
                                                                                          645
2041
2077
2240
2529
2788
1035
1047
                                                                                                                                                                                                                                                                     Local Similarity
nes 156; Conserv
                                                                                                                                                                                                                        1108
1120
1122
1129
1135
1139
                                                                                                                                                              1075
1077
1082
1086
1090
1093
                                                                                                                                                                                                           1101
                                                          CARBOHYD
CARBOHYD
                                                                       CARBOHYD
CARBOHYD
CARBOHYD
                                                                                         CARBOHYD
CARBOHYD
CARBOHYD
                                                                                                            CARBOHYD
CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                            SEQUENCE
                                                                                                                                                                                                                                                                                                                                                        Query Match
      CARBOHYD
              CARBOHYD
                   CARBOHYD
                           CARBOHYD
                                CARBOHYD
                                       CARBOHYD
                                                    CARBOHYD
                                                                                                                                                                           HELIX
TURN
STRAND
TURN
                                                                                                                                                                                                     STRAND
STRAND
TURN
                                                                                                                                                                                                                        STRAND
STRAND
STRAND
STRAND
                                                                                                                                                                                                                                                        STRAND
STRAND
HELIX
TURN
                                                                                                                                                                                                                                                                                  TURN
STRAND
TURN
STRAND
                                                                                                                                      HELIX
STRAND
STRAND
STRAND
                                                                                                                                                               TURN
                                                                                                                                                                                                                                                                                                                                     STRAND
                                                                                                                               STRAND
                                                                                                                                                                                                                                                                                                                  STRAND
                                                                                                                                                                                                                                                                                                                                                                Best Loca
Matches
                                                                                                                                                                                                                                                  TURN
                                                                                                                                                                                                                                                                                                                                                                                             qq
                                                                                                                                                                                                                                                                                                                                                                                                                      셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                 a
à
                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                  ò
```

```
between the Swiss Institute of Bioinformatics and the EMBL outstation—
the Buropean Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        This SWISS-PROT entry is copyright. It is produced through a collaboration
                                                  01-505 (Rel. 23, Created)
01-505 (Rel. 23, Last sequence update)
01-507 (Rel. 41, Last aenocation update)
62-67 (Rel. 41, Last aenocation update)
62-67 (Great and Contains: Capsid protein (Core protein) (P22);
62-68 (GP76) (RSI); Protein P7; Nonstructural protein NS2 (P21)
62-3-4-22-7; Protease/helicase NS3 (P70) (Hepacivitin)
62-3-4-21-98); Nonstructural protein NS54 (P4); Nonstructural protein NS54 (P5); Nonstructural protein NS54 (P6); Nonstructural protein NS58 (P6) (P0) (RNA-directed RNA polymerase) (EC 2.7.7.48)].
Hepatitis C virus (isolate Japanese) (HCV).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 -!- CATALYTIC ACTIVITY: Hydrolysis of four peptide bonds in the viral precursor polyprotein, commonly with Asp or Glu in the P6 position, Cys or Thr in Pl and Ser or Ala in Pl'.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          -1- SUBUNT: THE VIRION OF THIS VIRUS IS A NUCLECCAPSID COVERED BY LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS: PROTEIN M AND GLYCOPROTEIN E. THE NUCLECCAPSID IS A COMPLEX OF PROTEIN C AND MRNA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                              DISCUSSION OF SEQUENCE.

MEDLINE-91192160; PubMed-1849488;

Kato N., Hijkata M., Nakagawa M., Ootsuyama Y., Muraiso K.,

Ohkoshi S., Shimotohno K.;

"Molecular structure of the Japanese hepatitis C viral genome.";

FEBS Left. 280:325-328(1991).

-I- FUNCTION: THE SMALL PROTEINS NS2A, NS2B, NS4B ARE

HYDROPHOBIC, SUGGESTING A POSSIBLE MEMBRANE-RELATED FUNCTION.

NS3 AND NS5 MAY PLAY A ROLE IN THE VIRAL RNA REPLICATION.
                                                                                                                                                                                                                                                                                                                                                MEDLINE-91088550; PubMed-2175903;
Kato N., Hijikata M., Ootsuyama Y., Nakagawa M., Ohkoshi S., Sugimura T., Shimotohno K.;
"Molecular Cloning of the human hepatitis C virus genome from Japanese patients with non-A, non-B hepatitis.";
Proc. Natl. Acad. Sci. U.S.A. 87:9524-9528(1990).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     · I - SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY S29
                     PRT; 3010 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        HCV_capsid.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        HCV_env.
HCV_NS1.
HCV_NS3.
HCV_NS3.
HCV_NS4a.
HCV_NS4b.
HCV_NS5a.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       HCV_core
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      EMBL; D90208; BAA14233.1; ".
                     STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  InterPro; IPR001410;
InterPro; IPR002522;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         IPR002521;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          IPR002519;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              IPR002531;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         IPR002166;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     IPR000745;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          IPR002868;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       IPR001490;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       PIR; A39253; GNWVCJ
HSSP; P26663; 1JXP.
                                                                                                                                                                                                                                                                                                                                   SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                               Hepacivirus.
NCBL_TaxID=11116;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            MEROPS; S29.001;
MEROPS; U39.001;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          [RNA](N).
                     POLG_HCVJA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            InterPro;
InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        InterPro;
                                  P26662;
POLG_HCVJA
```

œ

```
1005 RRGKEILLGPADSFGEQCWRLLAPITAYSQQTRGLLGCIITSLTGRDKNQVDGEVQVLST 1064
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            54 ATQTFLATSINGVLWTVYHGAGTRTIASPKGPVTQMYTNVDKDLVGWQAPQGSRSLTPCT 113
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   53
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   3 KKGSVVIVGRINLSGD-----TAYAQQTRGEQGCQKTSHTGRDKNQVEGEVQIVST
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            9; Gaps
                                                                                                                                                                                       ProDom; PD186062; HCV_NS1; 1.
SNART; SM00487; DEXDc; 1.
SNART; SM00487; DEXDc; 1.
POlyprotein; Glycoprotein; Transferase; RNA-directed RNA polymerase;
Core protein; Glycoprotein; Envelope protein; Helicase; ATP-binding;
Transmembrane; Nonstructural protein; Hydrolase; Serine protease.
Transmembrane; Nonstructural protein; Hydrolase; Serine Protease.
                                                                                                                                                                                                                                                                                                                                                              RNA-DIRECTED RNA POLYMERASE (POTENTIAL)
                                                                                                                                                                                                                                                          CELLULAR AMINOPEPTIDASE.

CAPSID PROTEIN C (POTENTIAL).

MATRIX PROTEIN (POTENTIAL).

MAJOR ENVELOPE PROTEIN E (POTENTIAL).

NONSTRUCTURAL PROTEIN NSI (POTENTIAL).

NON-STRUCTURAL PROTEIN NS2 (POTENTIAL).
                                                                                                                                                                                                                                                                                                                            PROTEASE/HELICASE NS3 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4A (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4B (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                  CHARGE RELAY SYSTEM (BY SIMILARITY).
CHARGE RELAY SYSTEM (BY SIMILARITY).
CHARGE RELAY SYSTEM (BY SIMILARITY).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (POTENTIAL). (POTENTIAL). (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    80.7%; Score 823.5; DB 1; Length 3010; 75.5%; Pred. No. 2.1e-69; tive 23; Mismatches 18; Indels 9;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               MW: AA993794F46DB185
                                                                                                                                                                                                                                                                                                                                                                                                                                                     (GLCNAC.
(GLCNAC.
(GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                             (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       (GICNAC
                                                                                                                                                                                                                                                                                                                                                                                                                     ATP (POTENTIAL)
DECH BOX.
                                                                                                                                                                                                                                                                                                                                                                           POTENTIAL.
                                                                                                                                                                                                                                                                                                                                                                                                                                           N-LINKED
.erPro; IPR001650; Helicase_C.
.erPro; IPR07095; RNA_pol_DS_PS.
terPro; IPR007094; RNA_pol_DS_PS.
terPro; IPR007094; RNA_pol_DS_PS.
terPro; IPR007094; RNA_pol_DS_VIT.
tam; PF01542; HCV_corp; 1.
fam; PF01539; HCV_env; 1.
Pfam; PF01539; HCV_NS1; 1.
Pfam; PF01539; HCV_NS1; 1.
Pfam; PF01001; HCV_NS3; 1.
Pfam; PF01001; HCV_NS4x; 1.
Pfam; PF01001; HCV_NS4x; 1.
Pfam; PF01506; HCV_NS5x; 1.
Pfam; PF01506; HCV_NS5x; 1.
Pfam; PF00598; Viral_RGRP; 1.
ProDom; PD186062; HCV_NS1; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           STRGVAKAVDFIPVESLETIMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               327017
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Best Local Similarity 75.5
Matches 154; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         645
2041
2077
2240
2788
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       645 64
2041 204
2077 207
2240 224
2788 278
3010 AA;
                                                                                                                                                                                                                                                                                                                                                                                              1107
1165
11230
1316
196
209
234
250
305
                                                                                                                                                                                                                                                                                                                                                                                                         ACT_SITE
NP_BIND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                            CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                               ACT_SITE
                                                                                                                                                                                                                                                                                                                                                                                                                                                        CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                         TRANSMEM
                                                                                                                                                                                                                                                                                                                                                                                      ACT_SITE
                                                                                                                                                                                                                                                                                  CHAIN
CHAIN
CHAIN
CHAIN
                                                                                                                                                                                                                                                                                                                             CHAIN
                                                                                                                                                                                                                                                                       CHAIN
                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    õ
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMEL outstation the Buropean Bloinformatics Institute. There are not restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                       01-A0G-1992 (Rel. 23, Created)
01-A0G-1992 (Rel. 23, Last sequence update)
01-A0G-1992 (Rel. 24). Last sequence update)
01-A0G-1992 (Rel. 24). Last annotation update)
Genome polyprotein [Contains Capsid protein C (Core protein) (P22);
Envelope glycoprotein El (GP32) (GP35); Envelope glycoprotein E2
(GP68) (GP70) (NS1); Protein P7: Nonstructural protein NS2 (P21)
(EC 3.4.21.-98); Nonstructural protein NS4A (P4); Nonstructural protein NS4B (P27); Nonstructural protein NS5A (P56); Nonstructural protein NS5A (P66); Nonstructural NS5A (P66); Nonstructural protein NS5A (P66); Nonstructural protein NS5A (P66); Nonstructural protein NS5A (P66); Nonstructural NS5A (P66); NONSTRUCTURAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              [RNA](N).
-1- SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLECCAPSID COVERED BY LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS: PROTEIN M AND GLYCOROTEIN E. THE NUCLECCAPSID IS A COMPLEX OF PROTEIN C AND MRNA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SEQUENCE FROM N.A.
MEDLINE-92230232: PubMed-1314459;
Okamoto H., Kurah R., Okada S., I., Yamamoto K., Lizuka H., Tanaka 'Okamoto B., Tsuda F., Mishiro S.;
Flukuda S., Tsuda F., Mishiro S.;
"Full-length sequence of a hepatitis C virus genome having poor homology to reported isolates: comparative study of four distinct
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      -i- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY S29.
                                                                                                                                                  3033 AA
1185 CTRGVAKAVDFIPVESMETTMRSP 1208
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       InterPro; IPR001490; HCV_NS4b.
InterPro; IPR002868; HCV_NS5a.
InterPro; IPR0012165; HCV_RGAP.
InterPro; IPR007095; RNA_POL_DS_FS.
InterPro; IPR007094; RNA_POL_DS_FS.
InterPro; IPR007094; RNA_POL_PSvIr.
Pfam; PF01542; HCV_core; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     HCV_capsid.
HCV_core.
HCV_NS1.
HCV_NS2.
HCV_NS3.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  EMBL, D10988; BAA01761.1; -. PIR, A40250; GNWVJ8.
HSSP; P27958; 1HEI.
                                                                                                                                                  STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              IPRO02531; H
IPRO02518; B
IPRO04109; H
IPRO00745; H
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                InterPro; IPR001410;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            IPR002519;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              MEROPS; $29.001; -
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          NCBI_TaxID=11115;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Hepacivirus
                                                                                                                                             POLG_HCVJ8
P26661:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                InterPro;
InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            InterPro;
                                                                                                                        POLG_HCVJ8
                                                                                                                                                                          셤
```

6

```
pfam;
Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        79 IASPKGPVTQMYINVDKDLVGWQAPQGSRSLIPCTCGSSDLYLVTRHADVIPVRRRGDSR 138
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       19 TAYAQQTRGEQGCQKTSHTGRDKNQVEGEVQIVSTATQTFLATSINGVLWTVYHGAGTRT 78
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           139 GSLLSPRPISYLKGSSGCPLLCPACHAVGIFRAAVSTRGVAKAVDFIPVESLETTMRSP 197
                                                                                                                                                                                                                                                                                                            CELLULAR AMINOPEPTIDASE.
CAPSID PROTEIN C (POTENTIAL).
MATRIX PROTEIN C (POTENTIAL).
MATRIX PROTEIN C (POTENTIAL).
MAJOR ENVELOPE PROTEIN E (POTENTIAL).
NONSTRUCTURAL PROTEIN NSI (POTENTIAL).
NONSTRUCTURAL PROTEIN NSI (POTENTIAL).
NONSTRUCTURAL PROTEIN NSI (POTENTIAL).
NONSTRUCTURAL PROTEIN NSA (POTENTIAL).
NONSTRUCTURAL PROTEIN NSA (POTENTIAL).
NONSTRUCTURAL PROTEIN NSA (POTENTIAL).
NAMA DIRECTED RNA POLYMERASE (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              0; Gaps
                                                                                                                                                                                                                               Polyprotein; Glycoprotein; Transferase; RNA-directed RNA polymerase; Core protein; Coat protein; Envelope protein; Helicase; AFP-binding; Transmembrane; Nonstructural protein; Hydrolase; Serine protease. INTI_MET
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CHARGE RELAY SYSTEM (BY SIMILARITY).
CHARGE RELAY SYSTEM (BY SIMILARITY).
CHARGE RELAY SYSTEM (BY SIMILARITY).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ( POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ; Score 675; DB 1; Length 3033;
; Pred. No. 2.4e-55;
24; Mismatches 30; Indels (
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1A173E7E3381FD1A
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ATP (POTENTIAL).
DECH BOX.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   AA; 330177 MW;
                                                                     Pfam; PF02907; HCV_NS3; 1.
Pfam; PF01006; HCV_NS4s; 1.
Pfam; PF01001; HCV_NS4b; 1.
Pfam; PF01506; HCV_NS5a; 1.
Pfam; PF00998; Viral_RdRP; 1.
ProDom; P018606; HCV_NS1; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 66.2%;
69.8%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      st Local Similarity 69.83 tches 125; Conservative
                                                                                                                                                                                                            SM00487; DEXDC;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        3033
PF01539;
PF01560;
PF01538;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Query Match
Best Local S
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ACT_SITE
ACT_SITE
NP_BIND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       TRANSMEM
ACT_SITE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SEQUENCE
                                                                                                                                                                                      ProDom;
SMART; SI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CARBOHYD
                                                                                                                                                                                                                                                                                                                                                  CHAIN
                                                                                                                                                                                                                                                                                                                                                                                             CHAIN
CHAIN
CHAIN
CHAIN
CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               RESULT 8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           셤
    SOLUTION SOL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
```

PRT; 3033 AA

STANDARD;

POLG\_HCVJ6 ID POLG\_HCVJ6 AC P26660;

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss institute of Bioinformatics and the EMBL outstation the European Bioinformatics institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb.ch).
                                                                                                                                                                                                                                                                                                SEQUENCE FROM N.A.

MEDLINE-9204440; PubMed-1658196;
A MEDLINE-92044440; PubMed-1658196;
A MEDLINE-92044440; PubMed-1658196;
A Machida A., Miyakawa Y., Mayumi M.;
Mucleotide sequence of the genomic RNA of hepatitis C virus isolated T from a human carrier: comparison with reported isolates for conserved and divergent regions.";
J. Gen. Virol. 72:2697-2704(1991)
J. Gen. Virol. 72:2697-2704(1991)
J. Gen. Virol. 72:2697-2704(1991)
C -- FUNCTION: THE SNALL PROPEREDS MESPARE PREPARED FUNCTION.
NS3 AND NS5 MAY PLAY A ROLE IN THE VIRAL RNA REPLICATION.
NS3 AND NS5 MAY PLAY A ROLE IN THE VIRAL RNA REPLICATION.
C -- CATANTIC ACTIVITY: Hydrolysis of four peptide bonds in the viral precursor polyprotein, commonly with Asp or Glu in the P6 position, Cys or Thr in pl and Ser or Ala in Pl'.
C -- CATANTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
01-AUG-1992 (Rel. 23, Created)
01-AUG-1992 (Rel. 23, Last sequence update)
01-AUG-1992 (Rel. 21, Last annotation update)
03-FEB-2003 (Rel. 41, Last annotation update)
Genome polyprotein (Contains: Capsid protein C (Core protein) (P22);
Envelope glycoprotein E1 (GP32) (GP35); Envelope glycoprotein E2
(GP68) (GP70) (NR1); Protease/helicase R33 (P70) (Hepacivirin)
(EC 3.4.21.98); Nonstructural protein NS54 (P4); Nonstructural protein
NS48 (P27); Nonstructural protein NS54 (P4); Nonstructural protein
NS58 (P66) (P70) (RNA-directed RNA polymerase) (EC 2.7.7.48)}.
Hepatitis C virus (Isolate HC-76) (HCV).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     -1- SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS: PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C. AND MANA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY S29
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     INTERPRO) IPRO0252; HCV_capsid.
INTERPRO; IPRO0252; HCV_capsid.
INTERPRO; IPRO02519; HCV_core.
INTERPRO; IPRO02519; HCV_NS1.
INTERPRO; IPRO04519; HCV_NS2.
INTERPRO; IPRO04045; HCV_NS3.
INTERPRO; IPRO0445; HCV_NS4.
INTERPRO; IPRO0445; HCV_NS4.
INTERPRO; IPRO02668; HCV_NS5a.
INTERPRO; IPRO02668; HCV_NS5a.
INTERPRO; IPRO02669; HCV_NS5a.
INTERPRO; IPRO01669; HCV_NS5a.
INTERPRO; IPRO01695; RNA_POL_DS_PS.
INTERPRO; IPRO07095; RNA_POL_DS_PS.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Pfam; PF01543; HCV_caps1d; 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                EMBL; D00944; BAA00792.1; -. PIR; JQ1303; JQ1303.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         DEAD.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           HCV_core; 1
HCV_env; 1
HCV_NS1; 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              HCV_NS4b;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       InterPro; IPR001410;
InterPro; IPR002522;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          HSSP; P27958; 1HEI.
MEROPS; S29.001; -.
MEROPS; U39.001; -.
                                                                                                                                                                                                                                                                NCBI_TaxID-11113;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              PF01001;
PF01506;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        [RNA] (N).
                                                                                                                                                                                                                                              Hepacivirus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Pfam;
Pfam;
```

```
ô
                                                                                                                                                                                                                                                                                                                                                                                                                                        79 IASPKGPVTQMYTNVDKDLVGWQAPQGSRSLTPCTCGSSDLYLVTRHADVIPVRRGDSR 138
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                78
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     139 GSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVSTRGVAKAVDFIPVESLETTMRSP 197
                                       n: Envelope protein: Helicase: ATP-binding:
ral protein: Hydrolase: Serine protease.
RENVEDE FROM CAPEID FROTEIN C BY THE
CELLULAR AMINOPEPTIDASE.
CAPSID PROTEIN (POTENTIAL).
MATRIX PROTEIN (POTENTIAL).
MAJOR ENVELOPE PROTEIN E (POTENTIAL).
NONSTRUCTURAL PROTEIN NSI (POTENTIAL).
NONSTRUCTURAL PROTEIN NSI (POTENTIAL).
PROTEASE/HELICASE NSI (POTENTIAL).
NONSTRUCTURAL PROTEIN NSI (POTENTIAL).
NONSTRUCTURAL PROTEIN NSI (POTENTIAL).
NONSTRUCTURAL PROTEIN NSI (POTENTIAL).
NONSTRUCTURAL PROTEIN NSI (POTENTIAL).
POTENTIAL.
POTENTIAL.
                                                                                                                                                                                                                                                                                                                                                                                                                               19 TAYAQQTRGEQGCQKTSHTGRDKNQVEGEVQIVSTATQTFLATSINGVLWTVYHGAGTRT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Arabidopsis thaliana (Mouse-ear cress).
Eukaryota: Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
                                                                                                                                                                                                                                                                                                                                                                                                                Gaps
  Pfam; PF00999; Viral_RdRP; 1.
ProDom; PD186062; HCV_NS1 1.
SMART; SM00487; DEXDC; 1.
Polyprotein; Glycoprotein; Transferase: RNA-directed RNA polymerase; Core protein; Governotein; Envelope protein: Helicase: APP-binding; Transmembrane; Nonstructural protein; Hydrolase; Serine protease; INII_MET
                                                                                                                                                                      SIMILARITY).
SIMILARITY).
SIMILARITY).
                                                                                                                                                                                                                                                                                                                                                                                                               .;
0
                                                                                                                                                                                                                  (POTENTIAL)
                                                                                                                                                                                                                                  POTENTIAL
                                                                                                                                                                                                                                                     (POTENTIAL)
                                                                                                                                                                                                                                                                                               (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                 (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                          (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                   (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                   (POTENTIAL)
                                                                                                                                                                                                                                            (POTENTIAL
                                                                                                                                                                                                                                                                                                                (POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                             Length 3033;
                                                                                                                                                                                                                                                                                                                                                                                                              29; Indels
                                                                                                                                                                                                                                                                                                                                                                             F957F5C1A273BE9E CRC64;
                                                                                                                                                                       (BY
(BY
(BY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  HHOA_BRATH STANDARD, PKT; 321 AA. 09SEL7: 049507; 16-OCT-2001 (Rel. 40, Created) 16-OCT-2001 (Rel. 40, Last sequence update) 15-SEP-2003 (Rel. 42, Last annotation update) Protease HhoA, chloroplast precursor (EC 3.4.21..). HHOA OR AT4G18370 OR P28J12.30.
                                                                                                                                                                     CHARGE RELAY SYSTEM
CHARGE RELAY SYSTEM
CHARGE RELAY SYSTEM
                                                                                                                                                                                                                                                                                                                                                                                           ocore 673; DB 1; 1
Pred. No. 3.7e-55;
7; Mismatches 29:
                                                                                                                                                                                                                                                                                                                                                   (GLCNAC.
                                                                                                                                                                                                                                                                                                                                        (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                   (GLCNAC
                                                                                                                                                                                                ATP (POTENTIAL).
DECH BOX.
                                                                                                                                                                                                                                                                                                                                                                                             66.0%; Score 673;
68.7%; Pred. No. 3
                                                                                                                                                                                                                N.LINKED
                                                                                                                                                                                                                                                                                                                                                                  N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                             27;
Pfam: PF00271; helicase_C: 1.
                                                                                                                                                                                                                                                                                                                                                                           329165
                                                                                                                                                                                                                                                                                                                                                                                                      Local Similarity bs./
les 123; Conservative
                                                                                                                          1619
1866
2017
2017
369
1087
1111
11169
11241
1323
196
209
                                                                                                                                                                                                                                                                                                                                                                  2811
                                                                                                                                                                                                                                                                                                                                                                  2811 28:
3033 AA;
                                                                                                                                                          TRANSMEM
ACT_SITE
ACT_SITE
                                                                                                                                                                                     ACT_SITE
NP_BIND
SITE
                                                                                                                                                                                                                        CARBOHYD
                                                                                                                                                                                                                                                                                                             CARBOHYD
                                                                                                                                                                                                                                                                                                                                                CARBOHYD
CARBOHYD
                                                                                                                                                                                                                CARBOHYD
                                                                                                                                                                                                                                          CARBOHYD
                                                                                                                                                                                                                                                          CARBOHYD
                                                                                                                                                                                                                                                                     CARBOHYD
                                                                                                                                                                                                                                                                                     CARBOHYD
                                                                                                                                                                                                                                                                                                       CARBOHYD
                                                                                                                                                                                                                                                                                                                                 CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                           SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                             Query Match
                                                                                                                                                                                                                                                   CARBOHYD
                                                                                                                                                                                                                                                                                                                                        CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                   CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CHAIN
CHAIN
CHAIN
CHAIN
CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                       Best Loca
Matches
                                                                                                                                                    CHAIN
셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ð
```

```
WEBL_MACH_9702.

NERL_MACH_9702.

NERL_M
```

```
138
                                                                                                                                                                                                                                                                                                                                                                                                                                                               Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Sest Local
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  RESULT 11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              å
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ద
                     noved. Usage by and for commercial
(See http://www.lsb-sib.ch/announce/
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   12;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     140 ----SLLSPRPISYLK-----GSSGGPLLCPAGHAVGIFRAAVSTR--GVAKAV 182
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            182 --DPDNDLAVLKIETEGRELNPVVLGTSNDLRVGQSCFAI-----GNPYGYENTLTIGVV 234
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             235 SGLGREIPSPNGKSISEAIQTDADINSGNSGPLLDSYGHTIGVNTATFTRKGSGMSSGV 294
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      95 KTSPSVVYIEAIELPKTSSGDIL-----127
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               128 ----TGSGFVWDKLGHIVTNYHVIAKLATDQFGLQRCKVSLVDAKGTR--FSKEGKIVGL 181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              27
  as its content is in no
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           2 KKKGSVVIVGRINL - - - SCDTAYAQQTRGEQGCQKTSHTGRDKNQVEGEVQIVSTATQT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    58 FLATSINGVLW-----TVYH----TVYH-----GAGTRITASPKGPVTQM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            90 YINVDKDLVGWQAPQGSRSLTPCTCGSSDLYLVTRHADVIPVRRRGDSRG------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Pred. No. 0.66;

"w'ematches 68; Indels 100; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SEQUENCE FROM N.A.
SEQUENCE FROM N.A.
SEQUENCE FROM N.A.
SEQUENCE FROM N.A.
MEDLINE-20437337; PubMed-10984043;
Stover C.K., Pham X.-O.T., Erwin A.L., Mizoguchi S.D., Warrener P., Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M., Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M., Mizoguchi E.L., Golter S.N., Folger K.R., Was A., Larbig K., Lim R.H., Brody L.L., Spencer D.H., Wong G.K.-S., Wu Z., Paulisen II.T., Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulisen II.T., Relzer J., Saler M.H., Hancock R.E.W., Lory S., Olson M.V.; Complete genome sequence of Pseudomonas aeruginosa PAO1, an
                                                                                                                                                                                                                                                                                       Hydrolase; Serine protease; Chloroplast; Thylakoid; Transit peptide.
TRANSIT 26 CHLOROPLASI (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
Pseudomonadaceae; Pseudomonas
                                                                                                                                                                                                                                                                                                                                                                                     POLY-GLU.
CHARGE RELAX SYSTEM (POTENTIAL).
CHARGE RELAX SYSTEM (POTENTIAL).
CHARGE RELAY SYSTEM (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Mocal Similarity 21.3%; Score 87; DB 1; Length 321; Local Similarity 21.3%; Pred: No. 0.66; Length 321; Mismatch--
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  R -> G (IN REF. 1).
68DB81E0BD27A7A7 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    16-OCT-2001 (Rel. 40, Created)
16-OCT-2001 (Rel. 40, Last sequence update)
16-OCT-2001 (Rel. 40, Last annotation update)
Probable aromatic acid decarboxylase (EC 4.1.1.-).
              modified and this statement is not removed, motified and this statement is not removed, entities requires a license agreement (See hor send an email to license@isb.sib.ch),
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           209 AA.
                                                                                                                                                                                                                                                                                                                                            THYLAKOID.
PROTEASE HHOA.
non-profit institutions as long
                                                                                                         EMBL; AF114386; AAF24060.1; -.
EMBL; AL021710; CAA16717.1; ALT_SEO.
EMBL; AL161548; CAB78839.1; ALT_SEO.
HEROPS; SO1.279; -.
InterPro; IPR001940; Protease2C.
InterPro; IPR001940; Protease2C.
InterPro; IPR001954; Ser_protease_Try.
Pram; PF00089; trypsin; 1.
PRINTS; PR00834; PROTEASES2C.
                                                                                                                                                                                                                                                                                                                                                                                                          145 145 CH
186 186 CH
264 264 CH
40 40 RM3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                183 DF-IPVESLETTM 194
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          : | | | | :::: |:
NFAIPIDTVVRTV 307
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Pseudomonas aeruginosa.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          NCBI_TaxID-287;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           PAAD_PSEAE
                                                                                                                                                                                                                                                                                                                                                                                                          ACT_SITE
ACT_SITE
ACT_SITE
CONFLICT
                                                                                                                                                                                                                                                                                                                                            TRANSIT
CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Query Match
                                                                                                                                                                                                                                                                                                                                                                                          DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 29HX08;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               PA4019.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   PAAD_PSEAE
ID DT TO DT 
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          οy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         a
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bloinformatics and the EMBL outstation-the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              92 NVDKDLVGWQAPQGSRSLTP----CTCGSSDL-----YLVTRHADVIPVRRRGDS 137
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  RGSLLSPR--PIS----YLKGSSGGPLLCPAGHAVGIFRAAVSTRGVAKAVDFIPVES 189
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      43 QVEGEVQ-IVSTATQTFLATSINGVL------WTVYHGAGTRTIASPRGPVTQMYT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       29 QEEREVHFLISKAAQLVMATETDVALPAKPQAMQAFLTEYCGAAAGQI-----RVFG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        DEGI_ARATH STANDARD; PRT; 437 AA.

022609; Q9LK85;
16-OCT-2001 (Rel. 40, Created)
16-OCT-2001 (Rel. 41, Last sequence update)
28-FEB-2003 (Rel. 41, Last annotation update)
Protease Do-like 1, chloroplast precursor (EC 3.4.21..).
DEGPI OR DEGP OR AT3G27925 OR K16N12.18.
Arabidopsis thaliana (Mouse-ear cress).
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              STRAIN-CV. Columbia;
MEDLINE-20363099; PubMed-10907853;
Kanako T., Katoh T., Sato S., Nakamura A., Asamizu E., Tabata S.;
Ksructural analysis of Arabidopsis thaliana chromosome 3. II.
Sequence features of the 4,251,695 bp regions covered by 90 Pl, TAC
and BAC clones.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SEQUENCE FROM N.A., AND CHARACTERIZATION.
MEDLINE-98175982; PubMed-9507020;
Itzhaki H., Naveh L., Lindahl M., Cook M., Adam Z..;
Ildentification and characterization of Degp, a serine protease associated with the luminal side of the thylakoid membrane.";
J. Biol. Chem. 273:7094-7098(1998).
InterPro; IPR003382; Flavoprotein.
Pfam; PF02441; Flavoprotein; 1.
Hypothetical protein; Lyase; Decarboxylase; Complete proteome.
SEQUENCE 209 AA; 22367 MW; 01FD081CC495D3F6 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    DB 1; Length 209;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     61; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                8.4%; Score 85.5; DB 27.9%; Pred. No. 0.54; ative 16; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                         EMBL; AE004818; AAG07406.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  51; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    DNA Res. 7:217-221(2000).
                                                                                                                                                                                                                                                                                                                                                                                                                                                         PIR; H83144; H83144.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        [3]
SEQUENCE OF 104-118
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            190 LET 192
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         LNT 189
```

```
AAMP_HUMAN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      REPEAT
REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   EPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Best Loc
Matches
              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ð
                                                                                                                                                         This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bloinformatics and the FWBL outstation. the European Bloinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announce/or send an email to license@lsb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    7;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           102
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                209
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       103 PQGSRSLTPCTCGSSDLYLV------TRHADVIPVRRRGDSRGSLLSPRPI 147
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             210 PK--NKLRPIPVGVSADLLVGQKVFAIGNPFGLDHTLTTGVISGLRREIS--SAATGRPI 265
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            40-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                150 VPQGSGSGFVWDKQGHIVTNYHVIRGASDLRVTLADQTTFDAKVVGFDQDKDVAVLRIDA
         Kieselbach T., Bystedt M., Schroeder W.P.;
Submitted (JUL-2000) to the SWISS-PROT data bank.
--- FUNCATION: SERINE PROTEASE THAT IS REQUIRED AT HIGH TEMPERATURE.
MAY BE INVOLVED IN THE DEGRADATION OF DAWAGED PROTEINS. IN VIVO,
CAN DEGRADE BETA-CASEIN.
--- ENZYME REGULATION: INHIBITED BY PHENYLMETHYLSULFONYL FLUORIDE AND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                      CHARGE RELAY SYSTEM (POTENTIAL).
CHARGE RELAY SYSTEM (POTENTIAL).
CHARGE RELAY SYSTEM (POTENTIAL).
HSPPSSQLSNST -> SSTLFLHSPPSSHL (IN REF.
                                                                                                                                                                                                                                                                                                                                                                                      Transit peptide; Chloroplast; Thylakoid.
CHLOROPLAST (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SYL-----KGSSGGPLLCPAGHAVGIFRAAVSTRGVAKAVDF-IPVESL 190
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               266 QDVIQTDAAINPGNSGGPLLDSSGTLIGINTAIYSPSGASSGVGFSIPVDIV 317
                                                                                          -1 - SUBCELLULAR LOCATION: BOUND TO LUMINAL SIDE OF THE THYLAKOID
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    55;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             8.2%; Score 83.5; DB 1; Length 437;
26.2%; Pred. No. 2;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    55; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           -----TNVDKDLVGW-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1497B1AB3F5FF2A4 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  2).

V -> I (IN REF. 2).

P -> S (IN REF. 2).

G -> R (IN REF. 2).

G -> D (IN REF. 2).

L, -> HF (IN REF. 2).

L -> V (IN REF. 2).

I -> V (IN REF. 2).

I -> V (IN REF. 2).

O -> E (IN REF. 2).
                                                                                                                                                                                                                                                                                                                                                                                                                     PROTEASE DO-LIKE 1. SERINE PROTEASE.
                                                                                                              -1- INDICTION: By heat shock.
-1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S2C.
-1- SIMILARITY: Contains 1 PDZ/DHR domain.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Local Similarity 26.2%; Pred. No. 2; nes 45; Conservative 17; Mismatches
                                                                                                                                                                                                                                                                                           InterPro; IPR001478; PDZ.
InterPro; IPR001478; PDZ.
InterPro; IPR001940; Protease2C.
InterPro; IPR001254; Ser_protease_Try.
Pfam; PF00595; PDZ; 1.
Pfam; PF00689; trypsin; 1.
Pfam; PR0084; PR09810; 1.
PR0715; PR00844; PR07EASES2C.
SWART; SMO0228; PDZ; 1.
PR0SITE; PS50106; PDZ; 1.
                                                                                                                                                                                                                                                                                                                                                                                                              THYLAKOID.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         70 VYHGAGTRTIASPKGPVTQMY-----
                                                                                                                                                                                                                                                  EMBL; AF028842; AAC39436.1; -.
EMBL; AP000371; BAB02539.1; -.
EMBL; AP001302; BAB02539.1; JOINED.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        46213 MW;
                                                                                                                                                                                                                                                                                                                                                                                      iydrolase; Serine protease;
fRANSIT 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Columbia;
                                                                                 O-PHENANTHROLINE
                                                                                                                                                                                                                                                                                    MEROPS; S01.279;
                                                                                                                                                                                                                                                                                                                                                                                                                       104
152
171
201
280
                                                                                                       MEMBRANE
                                                                                                                                                                                                                                                                                                                                                                                                                                                     ACT_SITE
ACT_SITE
ACT_SITE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Query Match
                                                                                                                                                                                                                                                                                                                                                                                                            TRANSIT
CHAIN
DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                              NIAMO
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
ó
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ç
```

RESULT

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstationthe European Bioinformatics Institute. There are no restrictions on its mose by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@lsb-ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            197 WMEWH------PRAPVLLAGT-ADGNTWMWKVPNGDCKTFQGPNCPATCGR----- 240
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           123 TRHADVIPVRRR---GDSRGS-----LLSPRPISYLKGSSG--GPLLCPA------ 162
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           68 WIVYHGAGIRIIASPKGPVIQMYINVDKDLVGWQAPQGSRSL----TPCTCGSSDLYLV 122
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              241 -----VLPDGKRAVVGYEDGTIRIWDLKQGSPIHYLKGTEGHQGPLTCVAANQDGSLILT 295
                                                                                                                                                                                                                                                                                                                                                            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalla; Eutheria; Primates; Catarrhini; Hominidae; Homo.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               64.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ------GHAVGIFR----AAVSTRGVARAVDFIPVESL 190
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       HEPARIN-BINDING (POTENTIAL).
WD 1.
WD 2.
WD 3.
WD 4.
WD 5.
WD 6.
WD 6.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         8.1%; Score 83; DB 1; Length 452; 25.3%; Pred. No. 2.4; Live 13; Mismatches 47; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          DA1413D25EB236C0 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           GO; GO:0008201; F:heparin binding activity; TAS.
Interpor IPROUGL680; WT40.
Priam: PF00400; WD40; 8.
SMART; SM00320; WD40; 8.
PROSITE; PS00678; WD_REPEATS_1; 1.
PROSITE; PS50082; WD_REPEATS_2; 6.
PROSITE; PS50082; WD_REPEATS_2; 6.
                                           15-JUL-1998 (Rel. 36, Created)
15-JUL-1998 (Rel. 36, Last sequence update)
26-FEB-2003 (Rel. 41, Last annotation update)
Angio-associated migratory cell protein.
452 AA
                                                                                                                                                                                                                                                                                            SEQUENCE FROM N.A., AND PARTIAL SEQUENCE
PRT;
                                                                                                                                                                                                                                                                                                                      TISSUE-Brain;
MEDLINE-95262124; PubMed-7743515;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          49015 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              EMBL; M95627; AAA6889.1; -.
PIR; 139383; 139383.
Genew; HGNC:18; AAMP.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Conservative
STANDARD;
                                                                                                                                                                       Homo sapiens (Human)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          452 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Repeat; WD repeat.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Local Similarity tes 42; Conserv
                                                                                                                                                                                                                                             NCBI_TaxID-9606;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         603488;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ŏ
```

225 DMGEGYVVTTKHLYTKNGSSSAGPAQCPGGGGGG 258

296 GSVDCQAKLVSATTGKVVGVFRPETVASQPSLGEGEESESNSVESL 341

a

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       108
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          64. - NGVLWTVYHGAG---TRTIASPKGPVTQMYTNVDKDLVG------WQAPQGSRSLTPCT 113
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                109 GIGVYKHCVNGAGSSSTGITASPSTEICSQHAI----LVGGISKPFWLVPGGIGNNGNCG 164
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                165 CGGGGGGSSSSSSCIHIWLVPGGTGNNGNCGCGGGGGGSSSSSSCIHIKVENTDEOFL 224
                                                                                                                                                                                                                                                   Fraser C.M., Norris S.J., Weinstock G.M., White O., Sutton G.G., Dodson R., Gwinn M., Hickey E.K., Clayton R., Ketchum K.A., Sodergren E., Hardham J.M., McLeod M.P., Salzberg S., Peterson J., Khalak H., Richardson D., Howell J.K., Chidambaram M., Utterback T., McDonald L., Artiach P., Bowman C., Cotton M.D., Fujii C., Garland S., Hatch B., Horst K., Roberts K., Sandusky M., Weidman J., Smith H.O.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      54 AGSKLYATNGRLWEKELNGTGSWQKVSSSSVPTDSDK----KVMSIATDGNTFVLACVP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SGDIAYA - - - - - QQIRGEQGCQKTSH - - - - TGRDKNQVEGEVQIVSTATQIFLAISI -
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                114 C----GSSDLYLVTRHADVIP------VRRRGDSRGSLLSPRPISYLK-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  71; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                          Science 281:375-388(1998).
                                                                                                                                                   Spirochaetes; Spirochaetales; Spirochaetaceae; Treponema
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     POTENTIAL.
HYPOTHETICAL LIPOPROTEIN TP0136.
N-ACYL DIGLYCERIDE (POTENTIAL).
GLY/SER-RICH.
GLY/SER-RICH.
POLY-SER-RICH.
POLY-SER.
                                                                                                                                                                                                                                                                                                                                                                         "Complete genome sequence of Treponema pallidum, the syphilis spirochete.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                 -1- SIMILARITY: BELONGS TO THE TP013X FAMILY OF LIPOPROTEINS.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   DB 1; Length 485;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    77; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                C7A4CEEDC7DC5CED CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      TIGR; TP0136; -.
Hypothetical protein; Lipoprotein; Membrane; Signal;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     --GSSGGPLLCPAGHAVG 167
                                             16-OCT-2001 (Rel. 40, Created)
16-OCT-2001 (Rel. 40, Last sequence update)
16-OCT-2001 (Rel. 40, Last annotation update)
Hypothetical lipoprotein TP0136 precursor.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Score 78.5; Di
Pred. No. 6.9;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     EMBL; AE001199; AAC65137.1; ALT_INIT.
                                                                                                                                                                                                                       STRAIN-Nichols;
MEDLINE-98332770; PubMed-9665876;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    16;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                48984 MM;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  23.48;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                485 AA;
                                                                                                                                   Treponema pallidum
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Complete proteome. SIGNAL 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Similarity
                                                                                                                                                                                                       SEQUENCE FROM N.A.
                                                                                                                                                                    NCBI_TaxID-160;
                                                                                                                                                                                                                                                                                                                                                            J.C.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    50;
              Y136_TREPA
                                                                                                                                                     Bacteria;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                DOMAIN
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       E Com
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          LIPID
DOMAIN
DOMAIN
                                                                                                                                                                                                                                                                                                                                                            Venter
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 DOMAIN
Y136_TREPA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               q
                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                õ
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its most by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                   SEQUENCE FROM N.A.

MEDLINE-94102535; PubMed-7503814;

MEDGINE-94102535; DubMed-7503814;

Ramos R.G., Igloid G.L., Lichte B., Baumann U., Maier D.,

Schneider T., Brandstaetter J.H., Froehlich A., Fischbach K.-F.;

"The irregular chiasm C-roughest locus of Drosophila, which affects
aronal projections and programmed cell death, encodes a novel
immunoglobulin-like procein.";

Genes Dev. 7:2533-2547(1993).
                                                                                                                                                                                                                                                                                                                                                                           -:- SUBCELLULAR LOCATION: Type I membrane protein.
-:- TISSUE SPECIFICITY: POSTEMBRYONIC EXPRESSION IS STRONG IN THE DEVELOPING OPTIC LOBE AND IN THE EYE IMAGINAL DISC.
-:- DEVELOPMENTAL STAGE: STRONGLY EXPRESSED IN EMBRYOS. ALSO FOUND IN LATE LARVAL AND PUPAL STAGES.
-:- SIMILARITY: Contains 5 immunoglobulin-like C2-type domains.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          domain; Glycoprotein; Signal; Repeat;
                                                                                                                                                                                                                                                                                                                                -!- FUNCTION: REQUIRED FOR CORRECT AXONAL PATHWAY FORMATION IN THE OPTIC LOBE AND FOR PROGRAMMED CELL DEATH IN THE DEVELOPING
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         IRREGULAR CHIASM C-ROUGHEST PROTEIN.
EXTRACELLULAR (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (GLCNAC. .) (POTENTIAL). (GLCNAC. .) (POTENTIAL). (GLCNAC. .) (POTENTIAL). (GLCNAC. .) (POTENTIAL).
                                                                           Irregular chiasm C-roughest protein precursor (IRREC protein).
                                                                                               Drosophila melanogaster (Fruit fly),
Bukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CYTOPLASMIC (POTENTIAL).
IG-LIKE C2-TYPE 1.
IG-LIKE C2-TYPE 2.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    IG-LIKE C2-TYPE 3.
IG-LIKE C2-TYPE 4.
IG-LIKE C2-TYPE 4.
IG-LIKE C2-TYPE 5.
GLN-RICH (OPA-REPEAT).
N-LINKED (GLCNAC...) (
N-LINKED (GLCNAC...) (
N-LINKED (GLCNAC...) (
N-LINKED (GLCNAC...) (
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            EMBL; 221641; CAA79756.1; -.
EMBL; L11040; AAA16632.1; -.
PIR: A49448 A49448.
FINE A49468 A49448.
FINE A94468 A49448.
InterPro: IPR007110; Ig-like.
InterPro: IPR007110; Ig-like.
InterPro: IPR0073006; Ig_MRC.
Pfam; PF00047; Ig: 4.
SMART; SM00408; IG-2: 1.
PROSITE: PSS0835; IG_LIKE; 5.
Transmembrane; Immunoglobulin domain; Glycoprotein;
                            01-FEB-1995 (Rel. 31, Created)
01-FEB-1995 (Rel. 31, Last sequence update)
15-SEP-2003 (Rel. 42, Last annotation update)
764 AA.
 STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Cell adhesion
 CCR_DROME
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CHAIN
DOMAIN
TRANSMEM
DOMAIN
DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         OMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         DOMAIN
```

```
FT
                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                      a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       셤
                                                                                                                                                                                                                                                       õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the Buropean Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/
                                                                                                                                                            7:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                   109 --LIPCICGS--SDLYLVIRHADVIPVRRRGD----SRGSLLSPRPISYLKGSSGGPLLC 160
                                                                                                                                                                                                                                                                                 504 GKYNCTVVNDYGND----VAEIQLQAKKSVSLLMTIVGGISVVAFLLVLTILVVVYIKCK 559
                                                                                                                                                                                                                                                                                                                                                                                                      560 KRTKLPPADVISEHQITKNGGVSCKLEPGDRTSNYSDLKVDISGGYVPYGDYSTHYSPPP 619
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              620 QYLTTCSTKSNGSSTIMQNNHQNQLQLQQQQQQSHHQHHTQTTTLPMTFLTNSSGSL-- 677
                                                                                                                                                                                                                   27 GEQGCQKTSHTGRDKNQVEGEVQIVSTATQIFLATSING------VLWTVYHGAG 75
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               76 TRTIASP---------KGPVTQMYINVDKDLVGWQAPQGSRS-----
                                                                                                                                                               Gaps

    -!- PTM: SPECIFIC ENZYMATIC CLEAVAGES IN VIVO YIELD MATURE PROTEINS.
    -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY A2.

                                                                                                                                                        53;
507 507 N-LINKED (GLCNAC. . ) (POTENTIAL) 764 AA; 82947 MW; 262225D2B2A1C181 CRC64;
                                                                                           Length 764;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   01-MAY-1991 (Rel. 18, Created)
01-KKY-1991 (Rel. 18, Last sequence update)
01-KKY-1991 (Rel. 41, Last annotation update)
Pol polyprotein [Contains: Protease (EC 3 4.23.-); Reverse transcriptase (EC 2.7.7.49); Endonuclease).
                                                                                       7.7%; Score 78.5; DB 1; Length 76
20.3%; Pred. No. 12;
tive 27; Mismatches 69; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Viruses; Retroid viruses; Retroviridae; Gammaretrovirus NCBI_TaxiD=11840;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   PROSITE; PS00141; ASP_PROTEASE; 1.
PROSITE; PS50175; ASP_PROT_RETROV; 1.
Hydrolase; Transferase; RNA-directed DNA polymerase;
Aspartyl protease; Endonuclease; Polyprotein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            PRT; 1165 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               or send an email to license@isb-sib.ch)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  InterPro: IPR001995; Aspprotease_rtrv.
InterPro; IPR001969; Aspprotease_site.
InterPro; IPR002156; RNaseH.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SECTENCE FROM N.A.
MEDLINE-90051069; Pubmed-2683360;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          EMBL; M26927; AAA46810.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    InterPro; IPR001584; Rve.
InterPro; IPR000477; RVTse.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Gibbon ape leukemia virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Pfam; PF00075; rnaseH; l. Pfam; PF00655; rve; l. Pfam; PF00077; rve; l. Pfam; PF00078; rvt; l.
                                                                                              Ouery Match 7.78

Best Local Similarity 20.38

Matches 38; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   | :|
678 -TGSIIG 683
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 GNLJGL.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        161 PAGHAVG 167
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       PIR; B32595; GNLJGL
HSSP; P03355; IMML.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        MEROPS; A02.008;
   CARBOHYD
                                SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  POL_GALV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 RESULT 15
POL_GALV
                                                                                                                                                                                                                                                                                                                                                                                                  a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  WERNER PRESENTATION OF THE PRESENTATION OF THE PROPERTY OF THE
   5 5
                                                                                                                                                                                                                                                                       a
                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
```

```
411 PITPROVREFLGIAGFCRLWI----PGFASLAAPLYPLIKESIPFIWIEEHQQAFDHIKK 466
                                                                                                                                                                                                                                                                                                                                                                                                                         96 DLVGWQA---PQGSRSLTPCTCGSSDLYLVTRHADVIPVRRRGDSRGSL---LSP--RPI 147
                                                                                                                                                                                                                                                                                                                                                                                                                                                                           467 ALLSAPALALPDLTKPFT-----LYI-----DERAGVARGVLTQTLGPWRRPV 509
                                                                                                                                                                                                                ---RDKNQVEGEV 48
                                                                                                                                                                                                                                                                                                                     49 QIVSTATQTFLATSINGVLWTVYHGAGTRTIASPKGPVT-------OMYTNVDK 95
                                                                                                                                                              Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   148 SYLKG----SSGGPLLCPAGHAVGIFRAAVSTRGVAKAVDFIPVESLETTMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  64;
                                                                                                       Query Match 7.6%; Score 78; DB 1; Length 1165; Best Local Similarity 22.6%; Pred. No. 22; Matches 53; Conservative 33; Mismatches 85; Indels
                                                  1165 AA; 129886 MW; 8B7AFD54812B7E1A CRC64;
                                                                                                                                                                                                                3 KKGSVVIVGRINLSGDIAYAQQIRGEQGCQK-ISHIG-
PROTEASE.
BY SIMILARITY.
                         ACT_SITE
SEQUENCE
  CHAIN
```

Search completed: August 30, 2003, 19:13:49 Job time : 10.7567 secs

hepatitis

091rq8 h
091rq8 h
091rt1 h
091rt1 h
091rt1 h
091rt2 h
091rt2 h
091rt3 h
091rt3 h
091rt3 h
091rt3 h
091rt2 h
091rt2 h
091rt2 h
091rt2 h
001rt4 h
001rt4 h
001rt4 h
001rt4 h
001rt4 h
001rt5 h

Sequence:

on:

Searched:

hepatitis hepatitis

hepatitis hepatitis hepatitis hepatitis hepatitis hepatitis hepatitis hepatitis

hepatitis hepatitis hepatitis hepatitis hepatitis

hepatitis hepatitis hepatitis hepatitis

hepatitis hepatitis hepatitis

hepatitis hepatitis

```
SEQUENCE FROM N.A.
MEDLINE-20323484; PubMed-10864644;
Lal V.C., Zhong W., Skelton A., Ingravallo P., Vassilev V.,
Lal V.C., Anong W., Lau V.C.,
Bonis R.O., Hong Z., Lau J.Y.;
"Generation and characterization of a hepatitis C virus NS3 protease-dependent bovine viral diarrhea virus.";
[2]

1. Virol. 74:6339-6347(2000).
                                                                                                                                                                                                                                                                                                                       Genome polyprotein.
Mucosal disease virus.
Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                                                                       A Lai V.C.H., Hong 2.;
Submitted (MAY-2000) to the EMBL/GenBank/DDBJ databases.
Lai V.C.H., Hong 2.;
Submitted (MAY-2000) to the EMBL/GenBank/DDBJ databases.
RMSDP; P26663; 1JXP.
RMSDPS; S31.001; ...
RICEPTO; IPR001280; CDVir_endptseP80.
RICEPTO; IPR001410; DEAD.
RICEPTO; IPR001410; DEAD.
RICEPTO; IPR001410; DEAD.
RICEPTO; IPR001409; HCV_RGRP.
RICEPTO; IPR00105; Helicase_C.
RICEPTO; IPR00105; NAD_DNA_Dinding.
RICEPTO; IPR001058; RNA_POL_DS_PS.
RICEPTO; IPR001058; RNA_POL_DS_PS.
RICEPTO; IPR001059; RNA_POL_DS_PS.
RICEPTO; IPR001059; RNA_POL_PSVIr.
PFam; PF02907; HCV_NS3: 1.
PFam; PF02907; Holicase_C; 1.
                                                                                                                                                                                                                                                                                                 Created)
Last sequence update)
Last annotation update)
                                                                                                                                                                                                                                            ALIGNMENTS
091R81
091R08
091R00
036609
091R86
091R83
091R74
091R74
091R75
091R75
091R75
091R76
091R87
091R87
091R87
091R87
091R87
091R87
091R87
091R87
091R87
                                                                                                                                                                                                      Q91RR9
Q9DTE2
                                                                                                                                                                                                                                                                                                01-OCT-2000 (TrEMBLrel. 15, 01-OCT-2000 (TrEMBLrel. 15, 01-MAR-2003 (TrEMBLrel. 23,
SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                       NCBI_TaxID=11099;
838
837.5
837.5
837
836.5
                                                                                                                                                                                                                                                                                  Q9IFH8
                                                                                                                                                                                                                                                                   RESULT 1
Q9IFH8
Q91fhe mucosal dis
03679 hepatitis c
091756 hepatitis c
09165 hepatitis c
03463 hepatitis c
03668 hepatitis c
03668 hepatitis c
099wv9 hepatitis c
091rr8 hepatitis c
091rr5 hepatitis c
091rr5 hepatitis c
091rr5 hepatitis c
091rr5 hepatitis c
091rr7 hepatitis c
091rr7 hepatitis c
091rr7 hepatitis c
                                                   August 30, 2003, 19:00:22; Search time 37.5921 Seconds (without alignments) 1352.314 Million cell updates/sec
                                                                                                                                                                                                                                                                                                                                                                            Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.
                                                                                                 1 MKKKGSVVIVGRINLSGDTA.....VAKAVDFIPVESLETTMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                   Description
     GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.
                                                                                                                                                     Total number of hits satisfying chosen parameters:
                                                                                                                                        830525 seqs, 258052604 residues
                                                                                                                                                                                                                                                                                                                                                                                                            SUMMARIES
                                                                                                                                                                                           Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries
                                     - protein search, using sw model
                                                                                                                                                                                                                                                                                                                                                                                                                                                         036579
081756
091FE5
09ELS8
003463
036608
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               09PWU9
091RR8
091RR5
091RR5
091RR2
091RR3
                                                                                                                                                                                                                                                                                                                                                                                                                                                  Q9IFH8
                                                                                                                 BLOSUM62
Gapop 10.0 , Gapext 0.5
                                                                                                                                                                                                                                  sp_archea:*
sp_bacteria:*
sp_fungi:*
sp_human:*
sp_invertebrate:*
sp_mammal:*
sp_mhc:*
                                                                                                                                                                                                                                                                                                                                       sp_unclassified:*
                                                                                                                                                                                                                                                                                                                                               sp_rvirus:*
sp_bacteriap:*
sp_archeap:*
                                                                                                                                                                                                                                                                                                                              sp_vertebrate:*
                                                                                                                                                                                                                                                                                       sp_organelle:*
sp_phage:*
                                                                                                                                                                     Minimum DB seq length: 0
Maximum DB seq length: 200000000
                                                                                                                                                                                                                                                                                                                sp_rodent:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                US-09-965-594-20
1020
                                                                                                                                                                                                                                                                                                        sp_plant:*
                                                                                                                                                                                                                                                                                                                        sp_virus:*
                                                                                                                                                                                                                            SPTREMBL_23:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                 4040
3011
2436
3011
3011
3011
181
181
181
181
181
                                                                                                                                                                                                                                                                                                                                                                                                                                  Length
                                                                                                                                                                                                                                                                                                                                                                                                                          Query
                                                                                                                                                                                                                                                                                                                                                                                                                                                Title:
Perfect score:
                                                                                                                 Scoring table:
                                                                                                                                                                                                                                                                                                                                                                                                                                   Score
                                     OM protein
                                                                                                                                                                                                                            Database:
```

Result

```
PF01542;
PF01539;
PF01560;
                                                                                                                                                                                                                                                                                                                                                                                   Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      081756
081756;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  RESULT 3
Q81756
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ACCOCC ODER REPRESENTATION OF THE PROPERTY OF 
     ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           a
                                                                                                                                                                                                                                                    Ä
                                                                                                                                                                                                                                                                                                                                                                                     122 VTRHADVIPVRRGDSKGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVSTRGVAKA 181
                                                                                                                                                                                                                                                                                                                                                               62 SINGVLWTVYHGAGTRTIASPKGPVTQMYTNVDKDLVGWQAPQGSRSLTPCTCGSSDLYL 121
                                                                                                                                                                                                                                                                                                            S.
                                                                                                                                  ATP-binding; Helicase; Hydrolase; Nonstructural protein; Polyprotein; RNA-directed RNA polymerase; Transferase. SEQUENCE 4040 AA; 453073 MW; ADE87791D055B9DC CRC64;
                                                                                                                                                                                                                                                                                        5 GSVVIVGRINLSGD---TAYAQQTRGEQGCQKTSHTGRDKNQVEGEVQIVSTATQTFLAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              "Transmission of hepatitis C by intrahepatic inoculation with transcribed RNA.";

Cleance 277:570-574(1997).

-!- SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A LIPOPROTEIN ENVELOPE CONSISTS OF TWO PROTEINS:

PROTEIN M AND GIXCOPPOTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MRNA (BY SIMILARITY).

EMBL. AFURGOGO, AAB66324.1.

Interpro, IPROUI410, DEAD.
                                                                                                                                                                                                                                                  3; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
Hepacivirus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             MEDLINE-97373636; Pubmed-9228008;
Kolykhalov A.A., Agapov E.V., Blight K.J., Mihalik K., Feinstone
                                                                                                                                                                                                               DB 12; Length 4040;
                                                                                                                                                                                                           Ouery Match 85.5%; Score 872.5; DB 12; Length Best Local Similarity 89.2%; Pred. No. 1.4e-74; Matches 174; Conscrvative 5; Mismatches 13; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   01-JAN-1998 (TrEMBLrel. 05, Created)
01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    PRT; 3011 AA.
PRINTS; PR00729; CDVENDOPTASE.
SMART; SM00487; DEXDc; 1.
SMART; SM00480: HELICC; 1.
PROSITE; PS00037; MYB_1; 1.
PROSITE; PS50507; RDRP_POSITIVE; 1.
PROSITE; PS50501; RDRP_VIRAL; 1.
PROSITE; PS0531; RDRP_VIRAL; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Helicase_C.
RNA_pol_DS_PS.
RNA_pol_PSvir.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        HCV_capsid.
HCV_core.
HCV_env.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        HCV_NS4a.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   HCV_NS2.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                HCV_RdRP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         182 VDFIPVESLETIMRS 196
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           IPR002519; H
IPR002531; H
IPR002518; H
IPR004109; H
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        IPR002522;
IPR002521;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        IPR000745;
IPR001490;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   IPR007095;
IPR007094;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                IPR002166;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Genome polyprotein.
Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      NCBI_TaxID-11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   InterPro;
InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           STRAIN-H77
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Rice C.M.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        InterPro;
InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        InterPro;
InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   InterPro;
InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             InterPro;
                                                                                                                                                                                                               Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   036579;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  036579
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        a
                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                        à
                                                                                                                                                                                                                                                                                                                                                             à
                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                              셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               õ
```

```
113
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      114 CGSSDLYLVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAV 173
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            53
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      54 AIQTFLATSINGVLMTVYHGAGTRIIASPKGPVTQMYTNVDKDLVGWQAPQGSRSLTPCT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            3 KKGSVVIVGRIN -----LSGDTAYAQQTRGEQGCQKTSHTGRDKNQVEGEVQIVST
                                                                                                                                                                                                                         9; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae; Hepacivirus.
NCBI_TaxID=11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SEQUENCE FROM N.A.

Choo Q.-L., Richman K., Han J.;

Choo Q.-L., Richman K., Han J.;

The nucleotide sequence of the Hepatitis C viral genome.";

Submitted (MAY-1990) to the EMBL/GenBank/DDBJ databases.

EMBL: M32084; AA45677.1; -.

HSSP; P27958; JAIV.
                                                                                                                                                                                                                                                                                                                                                                                                                                   84.2%; Score 858.5; DB 12; Length
82.8%; Pred. No. 2.2e-73;
.ive 10; Mismatches 16; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     01-NOV-1996 (TrEMBLrel. 01, Created)
01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              PRT; 2436 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                174 STRGVAKAVDFIPVESLETTMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        InterPro; IPR001650; Helicase_C.
InterPro; IPR007095; RNA_pol_DS_PS.
InterPro; IPR007094; RNA_pol_DS_PS.
Pfam; PF01560; HCV_NS1; 1.
Pfam; PF01538; HCV_NS2; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      HSSF; K17308, 1A14.
InterPro; IPR001410; DEAD.
InterPro; IPR002531; HCV_NS1.
InterPro; IPR004109; HCV_NS3.
InterPro; IPR0004109; HCV_NS3.
InterPro; IPR001409; HCV_NS4b.
InterPro; IPR001409; HCV_NS4b.
InterPro; IPR001409; HCV_NS5b.
InterPro; IPR002166; HCV_NS5b.
                                    Pfam; PF01539; HCV_env; 1.
Pfam; PF01560; HCV_ENS; 1.
Pfam; PF01560; HCV_NS; 1.
Pfam; PF02907; HCV_NS; 1.
Pfam; PF01006; HCV_NS4x; 1.
Pfam; PF01001; HCV_NS4b; 1.
Pfam; PF01506; HCV_NS4b; 1.
Pfam; PF00171; Holicase_C; 1.
Pfam; PF00998; Viral_RGRP; 1.
ProDom; PD186062; HCV_NS1; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Genome polyprotein (Fragment). Hepatitis C virus.
HCV_capsid; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                          Best Local Similarity 82.8*
Matches 169; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              PRELIMINARY;
```

Ĥ

```
InterPro; IPR002166; HCV_RGRP.
InterPro; IPR001650; Helicase_C.
InterPro; IPR001095; Helicase_C.
InterPro; IPR001095; HRA_POL_DS_PS.
InterPro; IPR007094; RNA_POL_DS_PS.
InterPro; IPR007094; RNA_POL_DS_PS.
InterPro; IPR017094; RNA_POL_DS_PS.
InterPro; IPR01703; HCV_core; 1.
Pfam; PF01560; HCV_NS; 1.
Pfam; PF01560; HCV_NS; 1.
Pfam; PF01001; HCV_NS4z; 1.
Pfam; PF00271; Helicase_C: 1.
Pfam; PF00271; Helicase_C: 1.
Pfam; PF00266; HCV_NS1; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               174 STRGVAKAVDFIPVESLETIMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Genome polyprotein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   NCBI_TaxID-11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      STRAIN-colonel;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Hepatitis C
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Query Match
Best Local
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          O9ELS8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         RESULT 5
Q9ELS8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         DE PER PROPERTO DE PER PROPERT
                       셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ă
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         614
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       54 ATQTFLATSINGVLWTVYHGAGTRTIASPKGPVTQMYTNVDKDLVGWQAPQGSRSLTPCT 113
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            DR Pfam; PF01006; HCV NS4s; 1.

DR Pfam; PF01001; HCV NS4s; 1.

DR Pfam; PF01001; HCV NS4s; 1.

DR Pfam; PF01501; HCV NS4s; 1.

DR Pfam; PF05057; HCV NS2s; 1.

DR Pfam; PF00998; Viral_RdRP; 1.

DR Pfam; PF00998; Viral_RdRP; 1.

DR PF0009; PD180662; HCV NS1; 1.

DR PROSITE; PS50507; RDRP_POSITIVE; 1.

PROSITE; PS50501; RDRP_VIRAL; 1.

WATP-binding: Coat protein; Brvelope protein; Helicase; W Hydrolase; Monstructural protein; Polyprotein; Helicase; TRAN-directed RNA Polymerase; Transmembrane.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           23
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           3 KKGSVVIVGRIN-----I.SGDTAYAQOTRGEQGCQKTSHTGRDKNQVEGEVQIVST
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      555 RRGREILLGPADGMVSKGWRLLAPITAYAQQTRGLLGCIITSLTGRDKNQVEGEVQIVST
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            sequence..;
J. Gen. Virol. 82:1291-1297(2001).
J. Gen. Virol. THE VIRION OF THIS VIROLS CONSISTS OF TWO PROTEINS:
PROTEIN M AND GLYCOPROTEIN B. THE NUCLEOCAPSID IS A COMPLEX OF
PROTEIN C AND MRNA (BY SIMILARITY).
EMBL: A271632; AAF81759.1;
HSSP; P27958: 1AIV.
Interpro: IPRO000345, CytC.heme_bind.
Interpro: IPRO0003410; DEAD.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          9; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SEQUENCE FROM N.A. MEDIAGE 11369872; MEDIAME 2212: PubMed-11369872; Lanford R.E., Lee H., Chavez D., Guerra B., Brasky K.M.; Infectious cDNA clone of the hepatitis C virus genotype 1 prototype
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
Hepacivirus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Match 83.8%; Score 854.5; DB 12; Length 2436; Local Similarity 82.8%; Pred. No. 4e-73;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          17; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 2436 2436
2436 AA: 264734 MW; D7B9872900BE3125 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     01-0c7-2000 (TrEMBLrel. 15, Created)
01-0c7-2000 (TrEMBLrel. 15, Last sequence update)
01-0c7-2000 (TrEMBLrel. 13, Last annotation update)
Genome polyprotein.
Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              PRT; 3011 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       9; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    174 STRGVAKAVDFIPVESLETTMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      InterPro; IPR002522; HCV_capsid.
InterPro; IPR002521; HCV_core.
InterPro; IPR002531; HCV_env.
InterPro; IPR002531; HCV_NS1.
InterPro; IPR002518; HCV_NS3.
InterPro; IPR004109; HCV_NS3.
InterPro; IPR004109; HCV_NS3.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Matches 169; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  InterPro; IPR004109;
InterPro; IPR000745;
InterPro; IPR001490;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       NCBI_TaxID-11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Q9IFES
Q9IFES;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            2011

2017

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

2018

                    셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
```

```
1005 RRGREILLGPADGMYSRGWRLLAPITAYAQQTRGLLGCIITSLTGRDRNQVEGEVQIVST 1064
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        114 CGSSDLYLVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAV 173
                                                                                                                                                                                                                                                                                                                                                                                                                         54 ATQTFLATSINGVLWTVYHGAGTRTLASPKGPVTQMYTNVDKDLVGWQAPQGSRSLTPCT 113
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          3 KKGSVVIVGRIN-----LSGDTAYAQQTRGEQGCQKTSHTGRDKNQVEGEVQIVST
PROSITE; PS00487; DEXDC: 1.
PROSITE; PS00190; CYTCCHROME_C; 1.
PROSITE; PS50507; RDRP_POSITIVE; 1.
ATP-binding; Coat protein; Envelope protein; Glycoprotein; Helicase; Hydrolase; Nonstructural protein; Polyprotein;
RNA-directed RNA polymerase; Transferase; Transmembrane.
SEQUENCE 3011 AA; 327124 MW; 2489CE74AC864E58 CRC64;
                                                                                                                                                                                                                                                                           9; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 VIruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
Hepacivirus.
                                                                                                                                                                                                                          ch 83.8%; Score 854.5; DB 12; Length 3011; 1 Similarity 82.8%; Pred. No. 5.3e-73; 169; Conservative 9; Mismatches 17; Indels 9;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           01-MAR-2001 (TrEMBLrel. 16, Created)
01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
```

```
MEDIINE-92044440; PubMed-1658196; Okamoto H., Okada S., Sugiyama Y., Kurai K., Iizuka H., Machida A., Miyakawa Y., Mayumi M.;
Miyakawa Y., Mayumi M.;
"Nucleotide sequences of the genomic RNA of hepatitis C virus isolated from a human carrier: comparison with reported isolates for conserved and divergent regions.";
J. Gen. Virol. 72:2697-2704(1991).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      "Characterization and mapping of a B-cell immunogenic domain in hepatitis C virus E2 glycoprotein using a yeast peptide library."; virology 200:246-255(1994).
-1- SUBDMIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS: PROTEIN M AND GLYCOPROTEIN B. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MRNA (BY SIMILARITY).

EMBL: D10749; BAA0182.1; --
HSSP: P27958; 1HEI.
                                                                                                                                                                                                                                                                                                                                                                                                                STRAIN-HC-JI;
MEDLINE-93117120; PubMed-1335573;
Okamoto H., Kanai N., Mishiro S.;
Fruil-length nucleotide sequence of a Japanese hepatitis C virus isolate (HC-JI) with high homology to USA isolates.";
Nucleic Acids Res. 20:6410-6410(1992).
"The 5'-terminal sequence of the hepatitis C virus genome.";
Jpn. J. Exp. Med. 60:167-177(1990).
[2]
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ATP-binding; Coat protein; Envelope protein; Glycoprotein; Hydrolase; Nonstructural protein; Polyprotein; RNA-directed RNA polymerase; Transferase; Transmembrane. SEQUENCE 3011 AA; 327112 MW; 978905200250463B CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Mink M., Benichou S., Madaule P., Tiollais P., Prince A.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Okamoto H.;
Submitted (DEC-1992) to the EMBL/GenBank/DDBJ databases
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            InterPro; IPR007095; Helicase...
InterPro; IPR007095; RNA_pol_DS_PS.
InterPro; IPR007094; RNA_pol_PSvir.
Pfam; PF01543; HCV_capsid; 1.
Pfam; PF01542; HCV_core; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       PROSITE; PS50507; RDRP_POSITIVE; 1. PROSITE; PS50521; RDRP_VIRAL; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                MEDLINE-94174722; PubMed-7510436;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 InterPro; IPR002522; HCV_capsid.
InterPro; IPR002521; HCV_core.
InterPro; IPR002519; HCV_env.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  InterPro; IPR002519; RCV—Env.
InterPro; IPR002519; RCV MS1.
InterPro; IPR002518; HCV MS2.
InterPro; IPR0004169; HCV MS2.
InterPro; IPR001409; HCV MS4.
InterPro; IPR001490; HCV MS4a.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Pfam; PF01539; HCV_env; 1.
Pfam; PF01560; HCV_env; 1.
Pfam; PF01538; HCV_NS1; 1.
Pfam; PF01006; HCV_NS4; 1.
Pfam; PF01006; HCV_NS4; 1.
Pfam; PF01001; HCV_NS4; 1.
Pfam; PF01001; HCV_NS5; 1.
Pfam; PF010271; holicase_C; 1.
Pfam; PF00998; viral_RGRP; 1.
ProDom; P1186062; HCV_NS1; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         IPR002166; HCV_RdRP. IPR001650; Helicase_
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           InterPro; IPR001410; DEAD.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  DEXDC:
                                                                                                                             SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                          SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SMART; SM00487;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  STRAIN-HC-J1;
                                                                                                                                                     STRAIN-HC-J1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         inchauspe G.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Okamoto H
                                                   ä
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   54 ATQTFLATSINGVLWTVYHGAGTRIIASPKGPVTQMYTNVDKDLVGWQAPQGSRSLTPCT 113
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    114 CGSSDLYLVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAV 173
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     3 KRGSVVIVGRIN------LSGDTAYAQQTRGEQGCQKTSHTGRDKNQVEGEVQIVST
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       9; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           protein; Glycoprotein; Helicase;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
Hepacivirus.
NCBI_TaxID=11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      DB 12; Length 3011;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Okamoto H., Okada S., Sugiyama Y., Yotsumoto S., Tanaka T.,
Yoshizawa H.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                83.8%; Score 854.5; DB 12; Length
82.4%; Pred. No. 5.3e-73;
tive 11; Mismatches 16; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ATP-binding; Coat protein; Envelope protein; Glycoprotei Hydrolase; Monstructural protein; Polyprotein: RNA directed RNA polymerase; Transferase; Transmembane. SEQUENCE 3011 AA: 327107 HW; A6BECF5A3B3EE13F CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   01-NOV-1996 (TIEMBLEEL. 01, Created)
01-NOV-1996 (TIEMBLEEL. 01, Last sequence update)
01-MAR-2003 (TIEMBLEEL. 23, Last annotation update)
                                                                      | PR0021 | PR002515 | PR002515 | PR002515 | PR0002515 | PCV_NS2. | PR0002516 | HCV_NS2. | Pr0 | PR0004109 | HCV_NS3. | Pr0 | PR000409 | HCV_NS4. | Pr0 | PR000445 | HCV_NS4. | Pr0 | PR000266 | HCV_NS5a. | PR00266 | HCV_NS5a. | PR00266 | HCV_NS5a. | PR00266 | HCV_NS5a. | PR00266 | PCV_NS5a. | PR00266 | PCV_NS5a. | PR00266 | PCV_NS5a. | PR00266 | PCV_NS2 | PC0266 | PCV_NS3 |
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          PRI; 3011 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               174 STRGVAKAVDFIPVESLETTMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            PROSITE; PS00190; CYTOCHROME_C; 1. PROSITE; PS50507; RDRP_POSITIVE; 1. PROSITE; PS50521; RDRP_VIRAL; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 MEDLINE-91013116; PubMed-2170712;
                                                 HCV_capsid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Pfam; PF01506; HCV_NSSA; 1.
Pfam; PF00271; helicase_C: 1.
Pfam; PF00998; Viral_RdRP; 1.
Procom; P0186062; HCV_NS1; 1.
SMART; SM00487; DEXDC; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Best Local Similarity 82.48
Matches 168; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Genome polyprotein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SEQUENCE FROM N.A. STRAIN-HC-J1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Hepatitis C virus
                                                 InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Q03463
Q03463;
```

a

ò g ò q ö RESULT 6 Q03463

Helicase;

S

```
Hepacivirus
                                                                                                                                                                                                                SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    chimeras
                                                                                                                                                                                                                                                                     Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  09PWX5;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            O9PWX5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ω
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            RESULT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  O9PWX5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 HID DESCRIPTION OF THE PRESENCE OF THE PROPERTY OF THE PROPERT
a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ô
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        a
                                                                                                                                                                54 ATQTFLATSINGVLWTVYHGAGTRTIASPKGPVTQMYTNVDKDLVGWQAPQGSRSLTPCT 113
                                                                                                                                                                                                                                                                                                                                                 114 CGSSDLYLVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAV 173
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Yanagi M., Purcell R.H., Emerson S.U., Bukh J.,
"Transcripts from a single full-length cDNA clone of hepatitis C virus
are infectious when directly transfected into the liver of a
                                                                                                                                 3 KKGSVVIVGRIN-----LSGDTAYAQOTRGEQGCQKTSHTGRDKNQVEGEVQIVST
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ø
                                                                              Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Genome polyprotein.
Hepatitis C virus strain H77.
Viruses: ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Proc. Natl. Acad. Sci. U.S.A. 94:8738-8743(1997).
-1- SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY LIPOPROTEIN ENVELOPE. CONSIGTS OF TWO PROTEINS: PROTEIN MAND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MRNA (BY SIMILARITY).

EMBL: AFOll751: AAB67036.1; --
HSSP: P27958; IHEI.
                                                                           6
                   DB 12; Length 3011;
                                                                           Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  01-JAN-1998 (TrEMBLrel. 05, Created)
01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
                                                  6e-73;
                      Score 853.5; D
Pred. No. 6.6e-
8; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  PRT; 3011 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           InterPro; IPR001440; DEAD.
InterPro; IPR00252; HCV_capsid.
InterPro; IPR00252; HCV_capsid.
InterPro; IPR00252; HCV_care.
InterPro; IPR002531; HCV_core.
InterPro; IPR002513; HCV_NS1.
InterPro; IPR004109; HCV_NS2.
InterPro; IPR001409; HCV_NS4.
InterPro; IPR001409; HCV_NS4b.
InterPro; IPR001656; HCV_NS4b.
InterPro; IPR001666; HCV_NS4b.
InterPro; IPR001666; HCV_NS5b.
InterPro; IPR001669; HCV_NS5b.
InterPro; IPR001669; HCV_NS5b.
InterPro; IPR001699; RNA_pol_DS_PS.
InterPro; IPR001699; RNA_pol_DS_PS.
InterPro; IPR001699; RNA_pol_DS_PS.
InterPro; IPR01691; HCV_Capsid; IPR001699; HCV_CAPSID; HCV_CAPSID
                                                                                                                                                                                                                                                                                                                                                                                                                                                         174 STRGVAKAVDFIPVESLETTMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    MEDLINE-97385173; PubMed-9238047;
                      83.7%;
82.8%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               HCV_NS5a; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               HCV_NS4a;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           HCV_core;
                                                                         Matches 169; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  PRELIMINARY;
                                                  Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Hepacivirus.
NCBI_TaxID-63746;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   PF01001;
PF01506;
PF00271;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   PF02907;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               chimpanzee. ";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             STRAIN-H77
                         Query Match
                                                     Best Local
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  036608
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Pfam;
Pfam;
Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                        g
                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                      셤
                                                                                                                                                                                                                                                                                                                                              à
                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
```

```
;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    114 CGSSDLYLVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAV 173
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          23
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        3 KKGSVVIVGRIN-----LSGDTAYAQQTRGEQGCQKTSHTGRDKNQVEGEVQIVST
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                54 AIQTFLAISINGVLWTVYHGAGIRTIASPKGPVTQMYTNVDKDLVGWQAPQGSRSLIPCT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Gaps
                                                                                                            PROSITE; PSSO507; RDRP_POSITIVE; 1.
PROSITE; PSSO521; RDRP_VIRAL; 1.
ATP-binding; Coat protein; Bivelope protein; Glycoprotein; Helicase; Hydrolase; Monstructural protein; Polyprotein; Rna-directed RNA plymerase; Transferase; Transferase; Transferase; Transferase; Transferase; Changana, Alice Changana
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   MEDLINE-99420396; PubMed-10489358;
Yanagi M., Purcell R.H., Emerson S.U., Bukh J.;
Hepatitis C virus: an infectious molecular clone of a second major
genotype (2a) and lack of viability of intertypic la and 2a
chimeras.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                                            Length 3011;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           6
                                                                                                                                                                                                                                                                                                                                                                                                                                             ; Pred. No. 1e-72;
10; Mismatches 17; Indels
                                                                                                                                                                                                                                                                                                                        3011 AA; 327112 MW; 0B75E6B81CB5C198 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           01 WAY-2000 (TrEMBLrel. 13, Created)
01-WAY-2000 (TrEMBLrel. 13, Last sequence update)
01-WAR-2003 (TrEMBLrel. 23, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                       DB 12;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    PRT; 3015 AA
                                                                                                                                                                                                                                                                                                                                                                                                       83.5%; Score 851.5;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              174 STRGVAKAVDFIPVESLETTMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           HCV_caps1d.

HCV_core.

HCV_env.

HCV_NS1.

HCV_NS2.

HCV_NS3.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       HCV_NS4b.
HCV_NS5a.
HCV_RdRP.
Pfam; PF00998; Viral_RdRP; 1.
ProDom; PD186062; HCV_NS1; 1.
SMART; SM00487; DEXDC; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                             82.48;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Virology 262:250-263(1999).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      InterPro; IPR001410; DEAD.
InterPro; IPR002522; HCV_C
                                                                                                                                                                                                                                                                                                                                                                                                                                                 Best Local Similarity 82.4%
Matches 168; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         InterPro; IPR002521;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   IPR002519;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            InterPro; IPR002531;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              InterPro; IPR002518;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   InterPro; IPR004109;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      IPR002166;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           IPR001490;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                IPR002868;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Genome polyprotein.
Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  NCBI_TaxID-11103;
```

```
NCBI_TaxID=11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Hepacivirus
                                                                                                                                        InterPro;
InterPro;
InterPro;
                                                          InterPro;
                                                                                           InterPro;
                                                                                                                                                                           InterPro;
                                                                                                                                                                                                   InterPro;
                                                                                                                  InterPro;
                                                                                                                                                                                        InterPro;
                                                                                                                                                                                                           InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     091RR8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              RESULT 10
091RR8
AC 091RR8
DT 01-DEC
DT 01-DEC
DT NS3 PAI
DE NS3 PAI
OC HEPAL
OC HEPAL
OC HEPAL
OC HEPAL
OC HEPAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Matches
 õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                    113
                                                                                                                                                                                                                                                                                                                                                                                                                                 114 CGSSDLYLVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAV 173
                                                                                                                                                                                                                                                                                                                                        53
                                                                                                                                                                                                                                                                                                                                      3 KKGSVVIVGRIN-----LSGDTAYAQQTRGEQGCQKTSHTGRDKMQVEGEVQIVST
                                                                                                                                                                                                                                                                                                                                                                                   54 ATQTFLATSINGVLWTVYHGAGTRTIASPKGPVTQMYTNVDKDLVGWQAPQGSRSLTPCT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Bukh J.;
Submitted (AUG-1999) to the EMBL/GenBank/DDBJ databases.
-1- SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A LIFOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF
                                                                                                                                                                                PROSTE: MO0487; DEXDC: 1.
PROSTE: PS00192: DEXDC: 1.
PROSTE: PS00192: DDC GAD.HDC_YDC; 1.
PROSTE: PS50507: RDRP_POINTIVE; 1.
ATP-binding; Coat protein; Envelope protein; Glycoprotein; Helicase; Hydrolase; Nonstructural protein; Polyprotein;
RNA-directed RNA polymerase; Transferase; Transmembrane.
SEQUENCE 3015 AA; 328159 MM; B7D23BC1F190663A CRC64;
                                                                                                                                                                                                                                                                                                                9; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          a second major and 2a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
Hepacivirus
                                                                                                                                                                                                                                                                                          83.5%; Score 851.5; DB 12; Length 3015;
                                                                                                                                                                                                                                                                                                     ; Pred. No. 1e-72;
10; Mismatches 17: Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  MEDLINE-99420396; PubMed-10489358;
Yanaqi M., Purcell R.H., Emerson S.U., Bukh J.;
Yahaqatinis C virus: an infectious molecular clone of
genotype (2a) and lack of viability of intertypic la
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  01-MAY-2000 (TrEMBLrel. 13, Created)
01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              PRT; 3015 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       InterPro; IPR002129; PYITGOXAL GC.
InterPro; IPR007095; RNA_POL_DS_PS.
InterPro; IPR007095; RNA_POL_DS_PS.
InterPro; IPR007095; RNA_POL_DS_PS.
Pfam; PP01543; HCV_capsid: 1.
Pfam; PP01542; HCV_core; 1.
Pfam; PP01543; HCV_core; 1.
Pfam; PP01560; HCV_NS2; 1.
Pfam; PP01001; HCV_NS2; 1.
Pfam; PP01001; HCV_NS4; 1.
Pfam; PP01001; HCV_NS4; 1.
Pfam; PP01001; HCV_NS4; 1.
Pfam; PP00271; helicase 7.
Pfam; PP00998; Viral_RGRE; 1.
Prador, PP00998; Viral_RGRE; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                            174 STRGVAKAVDFIPVESLETTMRSP 197
InterPro; IPR001650; Helicase_C
                                                                                                                                                                                                                                                                                                     82.48;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Virology 262:250-263(1999)
                                                                                                                                                                                                                                                                                                   Best Local Similarity 82.4
Matches 168; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Genome polyprotein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                NCBI_TaxID-11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                chimeras.
                                                                                                                                                                                                                                                                                          Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             60Md60
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       RESULT 9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              a
                                                                                                                                                                                                                                                                                                                                                                                                       a
                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ö
                                                                                                                                                                                                                                                                                                                                        õ
                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  셤
```

```
114 CGSSDLYLVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAV 173
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             3 KKGSVVIVGRIN-----LSCDTAYAQQTRGEQGCQKTSHTGRDKNQVEGEVQIVST
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  9; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Helicase;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               NS3 protease (Fragment).
Hepatitis C virus.
Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             83.5%; Score 851.5; DB 12; Length 3015; 82.4%; Pred. No. 1e-72; Ive 10; Mismatches 17; Indels 9;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SMART: SM00487; DEXDC; 1.
PROSITE: PS00302; DDC_GAD_HDC_YDC; 1.
PROSITE: PS50507; RDRP_POSITIVE; 1.
PROSITE: PS50507; RDRP_VINIVE; 1.
ATP-binding; Coat protein; Envelope protein; Glycoprotein; Hydrolase; Nonstructural protein; Polyprotein;
RNA-directed RNA polymerase; Transmembrane.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ymerase; Transferase; Transmembrane.
328084 MW; E309F6318067D6CD CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 091RR8;
01-DEC-2001 (TrEMBLrel. 19, Created)
01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                181 AA
SIMILARITY).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            174 STRGVAKAVDFIPVESLETIMRSP 197
                                                                                                                                                                                                                                                                                          | IPR002868; HCV_NS5a.
| IPR00166; HCV_RGRP.
| IPR001650; Hellcase_C.
| IPR00129; Pyridoxal_deC.
| IPR007095; RNA_DOl_DS_PS.
                                                                                                                                                                                                                                                                                                                                                                                                                            Pfam; PP01543; HCV_capsid; 1. Pfam; PP01543; HCV_capsid; 1. Pfam; PP01542; HCV_care; 1. Pfam; PP01560; HCV_care; 1. Pfam; PP01560; HCV_NS2; 1. Pfam; PP01907; HCV_NS3; 1. Pfam; PP01006; HCV_NS4b; 1. Pfam; PP00101; HCV_NS4b; 1. Pfam; PP00271; helicase_C; 1. Pfam; PP00971; helicase_C; 1. Pfam; PP00998; Viral_RdRP; 1. ProDom; PD186662; HCV_NS1; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          82.48;
                                                                                      IPR001410; DEAD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             168; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                PRELIMINARY;
                                                                                                                                                                                                      IPR002518; H
IPR004109; H
IPR000745; H
IPR001490; H
                                                                                                                                                                              IPR002531;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SEQUENCE 3015 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Best Local Similarity
```

ä

~

```
79 IASPKGPVTQMYTNVDKDLVGWQAPQGSRSLTPCTCGSSDLYLVTRHADVIPVRRRGDSR 138
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              4 TAYAQOTRGLLGCIXTSLTGRDKNQVEGEVQIVSTAAQTFLATCINGVCWTVYHGAGTRT 63
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         139 GSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVSTRGVAKAVDFIPVESLETTMRS 196
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   124 GSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVCTRGVAKAVDFIPVENLETTMRS 181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              19 TAYAQQTRGEQGCQRTSHTGRDRNQVEGEVQIVSTATQTFLATSINGVLWTVYHGAGTRT
                                                                                                                                                                                                                                                                                                                                                                                             Holland-Staley C.A., Kovari L.C., Golenberg E., Mayers D.L.; "Genetic Diversity and response to IFN of the NS3 Protease Gene from Clinical Strains of the Hepatitis C Virus.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Holland-Staley C.A., Kovari L.C., Golenberg E., Mayers D.L.;
"Genetic Diversity and response to IFN of the NS3 Protease Gene from Clinical Strains of the Hepatitis C Virus.";
Submitted (APR-2001) to the EMBL/GenBank/DDBJ databases.
EMBL: AF369241; AAK54566.1; "
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
Hepacivirus.
                                                                                                                                                                                                                                                                 Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Length 181;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        11; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                            Submitted (APR-2001) to the EMBL/GenBank/DDBJ databases
EMBL; AR369238; AR54563.1; ...
InterPro; IRR04109; HCV_NS3.
Pfam; PF02907; HCV_NS3: 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  181 181
181 AA; 19123 MW; 1CAE817345ED809D CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            181 181
181 AA; 19084 MW; 3B5E8161F2100A72 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  01-DEC-2001 (TrEMBLrel. 19, Created)
01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
                                                                                                                                                  01-DEC-2001 (TrEMBLrel. 19, Created)
01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
01-DR-2003 (TrEMBLrel. 23, Last annotation update)
NS3 procease (Fragment).
Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Score 847; DB 12;
Pred. No. 7.2e-74;
                                                                                                                  Ş
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     2; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            83.0%;
92.7%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           InterPro; IPR004109; HCV_NS3.
Pfam; PF02907; HCV_NS3; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       01-DEC-2001 (TrEMBLrel. 19, 01-MAR-2003 (TrEMBLrel. 23, NS3 protease (Fragment). Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Matches 165; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              PRELIMINARY;
                                                                                                               PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                      SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             NCBI_TaxID-11103;
                                                                                                                                                                                                                                                                                                        NCBI_TaxID-11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             STRAIN-Pt. 4V;
                                                                                                                                                                                                                                                                                                                                                                               STRAIN-Pt.30;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Protease.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  NON_TER
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Protease.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SECUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Local
                                                                                                             091RR5
091RR5;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            091RR2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     091RR2
                                                                  RESULT 12
Q91RR5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                RESULT 13
Q91RR2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ద
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  64 IASPKGPVIQMYTNVDKDLVGMPAPQGSRSLTPCTCGSSDLYLVTRHADVIPVRRRGDSR 123
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         IASPKGPVTQMYTNVDKDLVGWQAPQGSRSLTPCTCGSSDLYLVTRHADVIPVRRRGDSR 138
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      79 IASPKGPVTQMYTNVDKDLVGWQAPQGSRSLTPCTCGSSDLYLVTRHADVIPVRRRGDSR 138
                                                                                                                                                                                                                                                                                                                                                                                                                      92
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      78
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         4 TAYAQOTRGLLGCIITSLTGRDKNQVEGEVQIVSTAAQTFLATCINGVCWTVYHGAGTRT 63
                                                                                                                                                                                                                                                                                                                                                                                                                                          4 TAYAQOTRGLLGCIITSLTGRDKNQVBGEVQIVSTAAQTFLATCINGVCWTVYHGAGTRT 63
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  GSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVSTRGVAKAVDFIPVESLETTMRS 196
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          139 GSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVSTRGVAKAVDFIPVESLETIMRS 196
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   19 TAYAQQTRGEQGCQKTSHTGRDKNQVEGEVQIVSTATQTFLATSINGVLWTVYHGAGTRT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  19 TAYAQQIRGEQGCQKISHIGRDKNQVEGEVQIVSIAIQIFLAISINGVLWTVYHGAGIRI
                                      FTRAIN-Pt.1x;

Holland-Staley C.A., Kovari L.C., Golenberg E., Mayers D.L.;

Holland-Staley C.A., Kovari L.C., Golenberg E., Mayers D.L.;

Genetic Diversity and response to IFN of the NS3 Protease Gene from Clinical Strains of the Hepatitis C Virus.";

Submitted (APR-2001) to the EMBL/GenBank/DDBJ databases.

EMBL; AF362235; AAK54560.1;

InterPro; IPR004109; HCV_NS3.

Pfam; PF02907; HCV_NS3.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            0; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                          ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Score 849; DB 12; Length 181;
Pred. No. 4.6e-74;
1; Mismatches 11; Indels (
                                                                                                                                                                                                                                                                                                                               Length 181;
                                                                                                                                                                                                                                                                                                                                                                          11; Indels
                                                                                                                                                                                                                                                               181
19130 MW; 85D91869299B7C35 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1 1
181 181
181 AA; 19130 MW; 85D91869299B7C35 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       01-DEC-2001 (TrEMBLrel. 19, Created)
01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
                                                                                                                                                                                                                                                                                                                             ; DB 12; 1
4.6e-74;
                                                                                                                                                                                                                                                                                                                             Score 849; DB pred. No. 4.6e-1; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 181 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 PRT;
                                                                                                                                                                                                                                                                                                                               83.2%;
93.3%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             83.2%;
93.3%;
                                                                                                                                                                                                                                                                                                                                                 Best Local Similarity 93.3
Matches 166; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Matches 166; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       NS3 protease (Fragment).
Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 PRELIMINARY;
                                                                                                                                                                                                                                                             [1]
SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            NCBI_TaxID=11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Hepacivirus
                                                                                                                                                                                                                                        NON_TER
NON_TER
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                NON_TER
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         29
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               139
                                                                                                                                                                                                                                                                                                                             Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Protease
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              NON_TER
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Local
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            091RT5
  ACCOCCOS ON THE REST OF THE RE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  요
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                   a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                      ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ö
```

78

Gaps

```
NCBI_TaxID-11103;
                                                                                                                                                                                                                                                                          Protease.
NON_TER
NON_TER
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                      Query Match
     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               음
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ö
                                                                           ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           79 IASPKGPVTQMYTNVDKDLVGWQAPQGSRSLTPCTCGSSDLYLVTRHADVIPVRRRGDSR 138
                                                                                                                                                                                                                         79 IASPKGPVTQMYTNVDKDLVGWQAPQGSRSLTPCTCGSSDLYLVTRHADVIPVRRRGDSR 138
                                                                                                                                                                                                                                                   64 IASPKGPVIQMYTNVDQDLVGWPAPQGARSLTPCTCGSSDLYLVTRHADVIPVRRRGDSR 123
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 78
                                                                                                                                                                          63
                                                                                                                                                                                                                                                                                                                      139 GSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVSTRGVAKAVDFIPVESLETTMRS 196
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             19 TAYAQQTRGEQGCQKTSHTGRDKNQVEGEVQIVSTATQTFLATSINGVLWTVYHGAGTRT
                                                                                                                                                        4 TAYAQQTRGLLGCIXISLIGRDKNQVEGEVQIVSTAXQTFLATCINGVCWTVYHGAGTRT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           STRAIN-Pt.1C;
Holland-Staley C.A., Kovari L.C., Golenberg E., Mayers D.L.;
Holland-Staley C.A., Kovari L.C., Golenberg E., Mayers D.L.;
"Genetic Diversity and response to The of the NS3 Protease Gene from Clinical Strains of the Hepatitis C Virus.";
Submitted (APR-2001) to the EMBL/GenBank/DDBJ databases.
EMBL; AF36214; AAK5439.1;
InterPro; IPRO04109; HCV_NS3.
                                                                             Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
Hepacivirus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
Hepacivirus.
                                                                           0;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               83.0%; Score 847; DB 12; Length 181; ilarity 92.7%; Pred. No. 7.2e-74; Conservative 2; Mismatches 11; Indels
                         Length 181;
                                                                           11; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         181 181
181 AA: 19114 MW; BEIDOB542F014E86 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Created)
Last sequence update)
Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             01-DEC-2001 (TTEMBLrel. 19, Created)
01-DEC-2001 (TTEMBLrel. 19, Last sequence update)
01-WAR-2003 (TTEMBLrel. 23, Last annotation update)
                       Score 847; DB 12;
Pred. No. 7.2e-74;
2; Mismatches 11;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         PRT; 181 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 181 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              091RR3;
01-DEC-2001 (TrEMBLrel. 19, Ct
01-DEC-2001 (TrEMBLrel. 19, Lz
01-MAR-2003 (TrEMBLrel. 23, Lz
NS3 procease (Frament).
Hepatitis C virus.
                         83.08;
92.78;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Pfam; PF02907; HCV_NS3; 1
                                                      Best Local Similarity 92.7
Matches 165; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        NS3 protease (Fragment). Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Local Similarity
nes 165; Conserva
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     NCBI_TaxID=11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  NON_TER
NON_TER
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Protease.
                         Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Q91RR3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Q91RT9
                                                                                                                                                                                                                                                                                                                                                                                                                                             RESOLT 14
091RT9
001RT9
DT 001-DE
DT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     RESULT 15
091RR3
1D Q91RR
AC Q91RR
DT 01-DE
DT 01-DB
DF NS3 P
DE NS3 P
OS Hepat
OC Virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        đ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            õ
                                                                                                                                                                   g
                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                         ογ
                                                                                                                                                                                                                                                                                                                                                                  8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
```

```
ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       79 IASPKGPVTQMYTNVDKDLVGWQAPQGSRSLTPCTCGSSDLYLVTRHADVIPVRRRGDSR 138
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     19 TAYAQQTRGEQGCQKTSHTGRDKNQVEGEVQIVSTATQTFLATSINGVLWTVYHGAGTRT 78
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  4 TAYAQQTRGLLGCIITSLTGRDKNQVEGEVQIVSTAAQTFLATCINGVCWIVYHGAGTRT 63
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        139 GSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVSTRGVAKAVDFIPVESLETTWRS 196
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      STRAIN-FL.4B;
Holland-Stelley C.A., Kovari L.C., Golenberg E., Mayers D.L.;
Holland-Stelley C.A., Kovari L.C., Golenberg E., Mayers D.L.;
Genetic Diversity and response to IFN of the NS3 Protease Gene from
Clinical Strains of the Hepatitis C Virus.";
Submitted (APR-2001) to the EMBL/GenBank/DDBJ databases.
EMBL; AF369240; AAK54565.1;
InterPro; IPR004109; HCV_NS3.
Pfam; PF02907; HCV_NS3; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                       Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                    ;
                                                                                                                                                                                                                                                                                                                                                                                            DB 12; Length 181;
                                                                                                                                                                                                                                                                                                                                                                                                                                                    Indels
                                                                                                                                                                                                                                                                              1 1
181 181
181 AA; 19115 MW; 5D85F88AD7AC1A11 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                         82.9%; Score 846; DB 12; 1
92.7%; Pred. No. 9e-74;
tive 2; Mismatches 11;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Search completed: August 30, 2003, 19:18:21 Job time : 38.5921 secs
                                                                                                                                                                                                                                                                                                                                                                                                                         Best Local Similarity 92.7% Matches 165; Conservative
SEQUENCE FROM N.A.
```

```
AR145272 Sequence
AR14526 Sequence
AR14526 Sequence
AR145273 Sequence
AR145273 Sequence
AR145273 Sequence
AR15275 Sequence
AR15259 Sequence
AR17810 Sequence
AR177810 Sequence
BD001911 Hepatitis
AR110828 Sequence
BD061910 Hepatitis
AR110828 Sequence
BD069982 Functiona
AF009606 Hepatitis
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ARI10831 Sequence
BD069985 Functiona
AR118686 Sequence
106348 Sequence 48
109328 Sequence 8
AR118692 Sequence
106440 Sequence 54
109329 Sequence 10
AR118696 Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       AR118696 Sequence
109331 Sequence 15
M32084 Hepatitis C
AR118703 Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  AF268278 Pestiviru
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          AR145257 Sequence
AR145252 Sequence
AR145265 Sequence
AR145270 Sequence
AR145274 Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             AR145253 Sequence
AR14526 Sequence
AR145261 Sequence
AR145269 Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    AR179057 Sequence
AR145264 Sequence
AR145268 Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               AR145262 Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                     Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ALIGNMENTS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SUMMARIES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  AR145255
AR145259
AR127810
BD081911
AR127809
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  BD081910
AR110828
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      BD069985
AR118686
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    AR110831
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    AR118692
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       AR118696
                                                                                                                                                                                                                                                                                                                                                               em_htgo_other:*
                                                                       em_htg_other:*
                                                                                                                                                                                                                                                                                           em_htgo_hum:*
em_htgo_mus:*
em_htg_hum:*
em_htg_inv:*
                                                                                                                                 em_htg_pln:*
                                                                                                                                                                                                  em_htg_mam:*
                                                                                                                                                                                                                                    em_htg_vrt:*
                                                                                                      em_htq_mus:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          DB
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      8157
8157
11932
9646
9646
9646
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
12980
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Query
Match Length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                651
1998
1998
2016
12734
651
651
1998
2016
648
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      892.5
8885.5
8881.5
8881.5
878.5
878.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      868
867.5
867.5
866.5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     864
863.5
863.5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       RESULT 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Š
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Command line parameters:
-WODEL-frame+p2n.model -DEV-xlp
-WODEL-frame+p2n.model -DEV-xlp
-WODEL-frame+p2n.model -DEV-xlp
-WOOPEL-frame+p2n.model -DEV-xlp
-WO-cgn2_1/GSPTO_spool/GS09965594/runat_29082003_151919_28310/app_query.fasta_1.2872
-DB-GenEmbl -OFMT-fastap -SUFFIX-spe -MINMATCH-0.1 -LOOPCL-0 -LOOPEXT-0
-UNITS-bblts -START-1 -END--1 -MATRIX-blosum62 -TRANS-human40.cd1 -LIST-45
-DOCALIGN-200 -THR_SCORE-pct -THR_MAX-100 -THR_MIN-0 -ALIGN-15 -MODE-LOCAL
-USER-GS09965594_eCGN_1 1_14686_8 runat_29082003_151919_28310 -NCPU-6 -ICPU-3
-NO_MARP -LARGEQUERY -NEG_SCORES-0 -WAIT -DSPBLOCK-100 -LONCLOG
-DEV_TIMBOUT-120 -WARN_TIMBOUT-30 -THRENDS-1 -XGAPOP-10 -XGAPEXT-0.5 -FGAPOP-6
                                                                                                                                                                                 August 30, 2003, 19:18:33 ; Search time 2560.57 Seconds (without alignments) 3147.423 Million cell updates/sec
                                                                                                                                                                                                                                                                                                                                                           1020
1 MKKKGSVVIVGRINLSGDTA.....VAKAVDFIPVESLETIMRSP
GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.
                                                                                                                                 OM protein - nucleic search, using frame_plus_p2n model
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       hits satisfying chosen parameters:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      2888711 seqs, 20454813386 residues
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries
                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Xgapop 10.0 , Xgapext
Ygapop 10.0 , Ygapext
Fgapop 6.0 , Fgapext
Delop 6.0 , Delext
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Minimum DB seq length: 0
Maximum DB seq length: 200000000
                                                                                                                                                                                                                                                                                                                             US-09-965-594-20
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              90.ba: *
90.ba: *
90.oa: *
90.ov: *
90.par: *
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   em_pat:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 em_pl:*
em_ro:*
em_sts:*
em_un:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               em_in:*
em_mu:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         em_fun:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 em_hum: *
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                em_om:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 *: vo_ma
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      em_or:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             GenEmbl: *
                                                                                                                                                                                                                                                                                                                                                                                                                                                           BLOSUM62
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   10...
11...
12...
12...
12...
12...
13...
14...
15...
16...
17...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
18...
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       oţ
                                                                                                                                                                                                                                                                                                                             Title:
Perfect score:
                                                                                                                                                                                                                                                                                                                                                                                                                                                           Scoring table:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Total number
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Database :
                                                                                                                                                                                                                                                                                                                                                                                          Sequence:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Searched:
                                                                                                                                                                                              Run on:
```

Sequence

Sequence

Sequence

```
AR145264
AR145264.1
                                                                                                                          Scores:
                                                                                                                                                                                                                                                                                                                                                102
                                                                                                                                                                                                                                                                                                                                                                                                        142
                                                                                                                                                                                                                                                                                                                                                                                                                                     162
                                                                                                                                                                                                                                                                                                                   82
                                                                                        source
                                                                                                    BASE COUNT
ORIGIN
                                                                                                                                  ..
No.:
                             ORGANISM
                                                  AUTHORS
TITLE
ACCESSION
VERSION
                                                                                                                          Alignment
                                           REFERENCE
                                                                       JOURNAL
              KEYWORDS
                                                                                FEATURES
                      SOURCE
                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                   à
                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                              õ
                                                                                                                                                                                                                                                                                                                                                             a
                                                                                                                                                                                                                                                                                                                                                                            ŏ
                                                                                                                                                                                                                                                                                                                                                                                        a
                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                      셤
                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                  ð
                                                                                                                                                                                                               g
                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                            9
                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                       ö
                                                                                                                                                                                                                                                                                                                                                                                                                     20-APR-2002
                                                                                                                                                                                                                                                                                                                                592
                                                                                                                                                                                                                                                                                                                                                                                          ValThrArgHisAlaAspValIleProValArgArgArgGlyAspSerArgGlySerLeu 141
                                                                                                                                                                                                                                            41
                                                                                                                                                                                                                                                                        61
                                                                                                                                                                                                                                                                                                     81
                                                                                                                                                                                                                                                                                SerIleAsnGlyValLeuTrpThrValTyrHisGlyAlaGlyThrArgThrIleAlaSer
                                                                                                                                                                                                                                                                                                              AsnGlnValGluGlyGluValGlnIleValSerThrAlaThrGlnThrPheLeuAlaThr
                                                                                                                                                                                                                                            AlaGlnGlnThrArgGlyGluGlnGlyCysGlnLysThrSerHisThrGlyArgAspLys
                                                         Unclassified.

1 (bases 1 to 12734)

Hong,Z., Lai,V.C.H. and Lau,J.Y.N.

Hepatitis C virus protease-dependent chimeric pestivirus

Patent: US 6326137-A 1 04-DEC-2001;

Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ValAspPheIleProValGluSerLeuGluThrThrMetArgSer 196
       linear
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Length:
Matches:
Conservative:
Mismatches:
Indels:
        DNA
   AR179057 12734 bp
Sequence 1 from patent US 6326137.
AR179057
                                                                                                                                                                                                 (1-12734)
                                                                                                          /organism="unknown"
1 2604 c 3295 g
                                                                                                                                                                                                  US-09-965-594-20 (1-197) x AR179057
                             AR179057.1 GI:20220612
                                                                                                                                              1.5e-65
892.50
92.82%
90.77%
87.50%
                                                                                                                                                              Percent Similarity:
Best Local Similarity:
                                                                                                                    ď
                                                                                                                   4032
                                            Unknown.
                                                    Unknown
                                                                                                                                         Alignment Scores:
Pred. No.:
                                                                                                                                                                                                                                                                                                                                                                            713
                                                                                                                                                                                                                                                                                                                                                                                                                                                                   893
                                                                                                                                                                                                                                             22
                                                                                                                                                                                                                                                                                      533
                                                                                                                                                                                                                                                                                                     62
                                                                                                                                                                                                                                                                                                                                                               102
                                                                                                                                                                                                                                                                                                                                                                                           122
                                                                                                                                                                                                                                                                                                                                                                                                                                                    162
                                                                                                                                                                                                                                                                         42
                                                                                                                                                                                                                                                                                                                   593
                                                                                                                                                                                                                                                                                                                                  82
                                                                                                                                                                                                                                                                                                                                               553
                                                                                                                                                                           Query Match:
DB:
                                                                                                      source
                     ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
                                                                                                                   BASE COUNT
ORIGIN
       LOCUS
DEFINITION
                                                                        AUTHORS
TITLE
JOURNAL
FEATURES
                                                                  REFERENCE
R179057
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              RESULT
                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                       q
                                                                                                                                                                                                                                                                                                                                                                           a
                                                                                                                                                                                                                                                                                                                                                                                        õ
                                                                                                                                                                                                                                                                                                                                                                                                       q
                                                                                                                                                                                                                                                                                                                                                                                                                      50
                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             셤
                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                       õ
                                                                                                                                                                                                                                                                                    a
                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                            οχ
```

08-AUG-2001

PAT

linear

AR145264 1998 bp D Sequence 105 from patent US 6211338.

DEFINITION

AR145264 LOCUS

```
08-AUG-2001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             AlaProGlnGlySerArgSerLeuThrProCysThrCysGlySerSerAspLeuTyrLeu 121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              LeuSerProArgProlleSerTyrLeuLysGlySerSerGlyGlyProLeuLeuCysPro 161
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     AlaGlyHisAlaValGlyIlePheArgAlaAlaValSerThrArgGlyValAlaLysAla 181
                                                                                                                                                                                                                                                                                                                                                                                                                                           61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ProLysGlyProValThrGlnMetTyrThrAsnValAspLysAspLeuValGlyTrpGln
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  544 TCGGGGCACGCTGTGGGCATCTTCCGGGCTGCCGTATGCACCCGGGGGGTTGCGAAGGCG
                                                                                                                                                                                                                                                                                                                                                                                                                            22 AlaGlnGlnThrArgGlyGluGlnGlyCysGlnLysThrSerHisThrGlyArgAspLy8
                                                                                                                                                                                                                                                                                                                                                                     GlySeryalvalllevalGlyArgIleAsnLeuSerGlyAsp.-----ThrAlaTyr
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   PAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            182 ValAspPheIleProValGluSerLeuGluThrThrMetArgSerPro 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1 (bases 1 to 1998).
Malcolm,B.A., Taremi,S.Shane., Weber,P.C. and Yao, Single-chain recombinant complexes of hepatitis C protease and NS4A cofactor peptide
Patent: US 6211338-A 105 03-APR-2001;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   linear
                                                                                                                                                                                                                                        1998
168
15
10
                                                                                                                                                                                                                                        Length:
Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 AR145268 1998 bp DNA Sequence 109 from patent US 6211338. AR142268 GI:15107135
                                                                                                                                                                                                                                                                                                                Gaps:
                                                                                                                                                                                                                                                                                                                                         US-09-965-594-20 (1-197) x AR145264 (1-1998)
                                                                                                                                                                 /organism="unknown"
595 c 569 a
                                                                                                                                           Location/Qualifiers
GI:15107131
                                                                                                                                                                                                                                       6.96e-66
885.50
93.37%
85.71%
86.81%
                                                                                                                                                        1998
                                                           Unclassified.
                                                                                                                                                                                                                                                                  Percent Similarity:
Best Local Similarity:
Query Match:
                    .
Unknown.
                                              Unknown
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              DEFINITION
ACCESSION
VERSION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      RESULT 3
AR145268
LOCUS
```

NS3

m

```
PAT 08-AUG-2001
                                                                                                                                                                                                                                                                                                                                                                                                   162 AlaGlyHisAlaValGlyIlePheArgAlaAlaValSerThrArgGlyValAlaLysAla 181
                                                                                                                                                                                                                                                                                                                                                           62 SerileAsnGlyValLeuTrpThrValTyrHisGlyAlaGlyThrArgThrIleAlaSer 81 ::::!!{|||||||
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 244 TGCGTCAACGGCGTGTGTTGGACCGTTTACCATGGTGCTGGCTCAAAGACCTTAGCCGGC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           142 LeuSerProArgProlleSerTyrLeuLysGlySerSerGlyGlyProLeuLeuCysPro
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     42 AsnGlnValGluGlyGluValGlnIleValSerThrAlaThrGlnThrPheLeuAlaThr
                                                                                                                                                                                                                                                                             22 AlaGlnGlnThrArgGlyGluGlnGlyCysGlnLysThrSerHisThrGlyArgAspLys
                     1 (bases 1 to 1998)
Malcolm, B.A., Taremi, S.Shane., Weber, P.C. and Yao, N.
Single-chain recombinant complexes of hepatitis C virus
profease and NS4A cofactor peptide
Patent: US 6211338-A 104 03-APR-2001;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          182 ValAspPhelleProValGluSerLeuGluThrThrMetArgSerPro 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      linear
                                                                                                                                             1998
168
14
11
                                                                                                                                             Length:
Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        DNA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     AR145254 651 bp
Sequence 95 from patent US 6211338.
AR145254
                                                                                                      424
                                                                                                                                                                                               Gaps:
                                                                                                                                                                                                                   US-09-965-594-20 (1-197) x AR145263 (1-1998)
                                                                                         /organism~"unknown"
596 c 568 q
                                                                         Location/Qualiflers
                                                                                                                                             1.52e-65
881.50
92.86%
85.71%
86.42%
                                                                                    1998
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Unknown.
Unclassified
            Unclassified
                                                                                                                                                                Percent Similarity:
Best Local Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    AR145254.1
                                                                                                     410 a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Unknown.
                                                                                                                                     Scores:
                                                                                                                                                                                      Query Match:
                                                                                    source
                                                                                                                                              Pred. No.:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               DEFINITION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ORGANISM
   ORGANISM
                                                                                                      BASE COUNT
                                AUTHORS
TITLE
                                                                                                                                    Alignment
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ACCESSION
                        REFERENCE
                                                              JOURNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   VERSION
KEYWORDS
SOURCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   RESULT 5
AR145254
                                                                         FEATURES
                                                                                                                  ORIGIN
                                                                                                                                                                                                                                       ô
                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                               d
                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                               a
                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                      a
                                                                                                                                                                                                                                                                                                                                                                                                                                            ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                               8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               合
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           PAT 08-AUG-2001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             363
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ValThrArgHisAlaAspVallleProValArgArgArgGlyAspSerArgGlySerLeu 141
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         483
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    243
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  423
                                                                                                                                                                                                                                                                                                                                                                                                    303
                                                                                                                                                                                                                                                                                                                                                                                                                                                               AlaProGlnGlySerArgSerLeuThrProCysThrCysGlySerSerAspLeuTyrLeu 121
                                                                                                                                                                                                                                                                                                                                                                                                                      ProLysGlyProValThrGlnMetTyrThrAsnValAspLysAspLeuValGlyTrpGln 101
                                                                                                                                                                                                                                                                                                 41
                                                                                                                                                                                                                                                                                                                                        61
                                                                                                                                                                                                                                                                                                                                                                               81
                                                                                                                                                                                                                                                                                                                                                                                        AsnGlnValGluGlyGluValGlnIleValSerThrAlaThrGlnThrPheLeuAlaThr
                                                                                                                                                                                                                                                                                                                                                                            SerIleAsnGlyValLeuTrpThrValTyrHisGlyAlaGlyThrArgThr1leAlaSer
                                                                                                                                                                                                                                                                                               22 AlaGlnGlnThrArgGlyGluGlnGlyCysGlnLysThrSerHisThrGlyArgAspLys
                                                              virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ValAspPheIleProValGluSerLeuGluThrThrMetArgSerPro 197
                                        1 (bases 1 to 1998)
Malcolm, B.A., Taremi, S.Shane., Weber, P.C. and Yao, N.
Single-chain recombinant complexes of hepatitis C vi
protease and Ns4A cofactor peptide
Patent: US 6211338-A 109 03-APR-2001;
Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        linear
                                                                                                                                                                1998
167
16
10
                                                                                                                                                                          Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1998 bp I
from patent US 6211338.
                                                                                                                        423
                                                                                                                                                                Length:
                                                                                                                                                                                                                                     US-09-965-594-20 (1-197) x AR145268 (1-1998)
                                                                                                              /organism="unknown"
595 c 569 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Sequence 104 from paten
AR145263
AR145263.1 GI:15107130
                                                                                                                                                               1.25e-65
882.50
93.37%
85.20%
86.52%
                                                                                                      1. .1998
                                Unclassified
                                                                                                                                                                                   Percent Similarity:
Best Local Similarity:
                                                                                                                        ಥ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          AR145263
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Unknown
         Unknown
                     Unknown
                                                                                                                                                      Alignment Scores:
                                                                                                                                                                                                                                                                                                                                       42
                                                                                                                                                                                                                                                                                                                                                                               62
                                                                                                                                                                                                                                                                                                                                                                                                                      82
                                                                                                                                                                                                                                                                                                                                                                                                                                                              102
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       124
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                184
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     122
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   RESULT 4
AR145263
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
                                                                                                                                                                                                        Query Match:
                                                                                                      source
                                                                                                                                                                Pred. No.:
                      ORGANISM
                                                                                                                        BASE COUNT
                                        REFERENCE
AUTHORS
TITLE
                                                                                 JOURNAL
KEYWORDS
                                                                                           FEATURES
                                                                                                                                   ORIGIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           q
                                                                                                                                                                                                                                                                           q
                                                                                                                                                                                                                                                                                                å
                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                      a
                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                         셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ð
                                                                                                                                                                                                                                                         ò
```

243

61

4

303

141 483 161 543 603

us-09-965-594-20.rge

```
Unclassified
                                                                                                                   Percent Similarity:
Best Local Similarity:
                                                            Ø
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Unknown.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Unknown
                                                                                       Alignment Scores:
                                                                                                                                                                                                                                                                                                      62
                                                                                                                                                                                                                                                                                                                                            83
                                                                                                                                                                                                                                                                                                                                                              304
                                                                                                                                                                                                                                                                                                                                                                                102
                                                                                                                                                                                                                                                                                                                                                                                                   364
                                                                                                                                                                                                                                                                                                                                                                                                                      122
                                                                                                                                                                                                                                                                                                                                                                                                                                       424
                                                                                                                                                                                                                                                                                                                                                                                                                                                           142
                                                                                                                                                                                                                                                                                                                                                                                                                                                                              484
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   544
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 162
                                                                                                                                      Query Match:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         DEFINITION
ACCESSION
                                         source
                                                                                              Pred. No.:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ORGANISM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            REFERENCE
AUTHORS
TITLE
                                                          BASE COUNT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            VERSION
KEYWORDS
SOURCE
                     JOURNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             RESULT 7
AR145262
                               FEATURES
   TITLE
                                                                    ORIGIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                q
                                                                                                                                                                                                      g
                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                             ద
                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                    q
                                                                                                                                                                                                                                                                                                                                          à
                                                                                                                                                                                                                                                                                                                                                            요
                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                               셤
                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   PAT 08-AUG-2001
                                                                                                                                                                                                                                                                                                                                                           AlaProGlnGlySerArgSerLeuThrProCysThrCysGlySerSerAspLeuTyrLeu 121
                                                                                                                                                                                                                                                                                                                                                                                                                                        141
                                                                                                                                                                                                                                                                                                                                                                                                                                                  303
                                                                                                                                                                                                                                                                                                                                                                                                                      GCGCCCCCCGGGGGCGCGTTCCTTGACACCATGCACCTGTGGCAGCTCAGACCTTACTTG 423
                                                                                                                                                                                                                         64 GGTTCTGTTGTTGTTGGTAGAATTATTTTATCTGGTAGTGGTAGTATCACGGCCTAC 123
                                                                                                                                                                                                                                                                                  61
                                                                                                                                                                                                      GlyServalvalileValGlyArgileAsnLeuSerGlyAsp------ThrAlaTyr 21
                                                                                                                                                                                                                                             AlaGlnGlnThrArgGlyGluGlnGlyCysGlnLysThrSerHisThrGlyArgAspLys 41
                                                                                                                                                                                                                                                                                                                        SerIleAsnGlyValLeuTrpThrValTyrHisGlyAlaGlyThrArgThrIleAlaSer 81
                                                                                                                                                                                                                                                                                                                                          TGCGTCAACGGCGTGTTTGGACCGTTTACCATGGTGCTGGCTCAAAGACCTTAGCCGGC
                                                                                                                                                                                                                                                                                                                                                                                                                                      ValThrArgHisAlaAspValIleProValArgArgArgGlyAspSerArgGlySerLeu
                                                                                                                                                                                                                                                                                 AsnGlnValGluGlyGluValGlnIleValSerThrAlaThrGlnThrPheLeuAlaThr
1 (bases 1 to 651)
Malcolm,B.A., Taremi,S.Shane., Weber,P.C. and Yao,N.
Signle-chain recombinant complexes of hepatitis C virus NS3
protease and NS4A cofactor peptide
Patent: US 6211338-A 95 03-APR-2001;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1 (bases 1 to 1998)
Malcolm, B.A., Taremi, S.Shane., Weber, P.C. and Yao, N.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             648
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         182 ValAspPheIleProValGluSerLeuGluThrThrMetArgScr 196
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    linear
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  651
167
15
10
                                                                                                                  Length:
Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    DNA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Sequence 108 from patent US 6211338.
AR145267
AR145267.1 GI:15107134
                                                                             144
                                                                                                                                                                  Gaps:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1998 bp
                                                                                                                                                                                     US-09-965-594-20 (1-197) x AR145254 (1-651)
                                                                 /organism="unknown"
187 c 200 g
                                                Location/Qualifiers
                                                                                                                 7.52e-66
878.50
93.33%
85.64%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Unknown.
Unclassified.
                                                                                                                                   Percent Similarity:
Best Local Similarity:
Query Match:
                                                                             Ø
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   AR145267
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Unknown.
                                                                                                         Alignment Scores:
Pred. No.:
                                                                                                                                                                                                                                             22
                                                                                                                                                                                                                                                                                                                      62
                                                                                                                                                                                                                                                                                                                                            244
                                                                                                                                                                                                                                                                                                                                                             82
                                                                                                                                                                                                                                                                                                                                                                               304
                                                                                                                                                                                                                                                                                                                                                                                                   102
                                                                                                                                                                                                                                                                                                                                                                                                                    364
                                                                                                                                                                                                                                                                                                                                                                                                                                                          424
                                                                                                                                                                                                                                                                                 42
                                                                                                                                                                                                                                                                                                                                                                                                                                       122
                                                            source
                                                                            BASE COUNT
ORIGIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           DEFINITION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ORGANISM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ACCESSION
VERSION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               REFERENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          AUTHORS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             RESULT 6
AR145267
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         KEYWORDS
SOURCE
            AUTHORS
                                       JOURNAL
  REFERENCE
                                                 FEATURES
                     TITLE
                                                                                                                             Score:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  LOCUS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ö
                                                                                                                                                                                                      ò
                                                                                                                                                                                                                         ద
                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                              a
                                                                                                                                                                                                                                                                                 ö
                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                         q
                                                                                                                                                                                                                                                                                                                                                             á
                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                    셤
                                                                                                                                                                                                                                                                                                                                                                                                                                       ç
                                                                                                                                                                                                                                                                                                                                                                                                                                                        q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                             õ
```

```
PAT 08-AUG-2001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ProLysGlyProValThrGlnMetTyrThrAsnValAspLysAspLeuValGlyTrpGln 101
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        LeuSerProArgProlleSerTyrLeuLysGlySerSerGlyGlyProLeuLeuCysPro 161
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         543
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       64 GGTTCTGTTATTGTTGGTAGAATTATTTATCTGGTAGTGGTAGTATCACGCCTAC 123
                                                                                                                                                                                                                                                                                                                                                                                                      CCAAAGGGGCCAATCACCCAGATGTACACTAATGTGGACCAGGACCTCGTCGGCTGGCAG 363
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          AlaProGlnGlySerArgSerLeuThrProCysThrCysGlySerSerAspLeuTyrLeu 121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ValThrArgHisAlaAspValIleProValArgArgArgGlyAspSerArgGlySerLeu 141
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            81
                                                                                                                                                                                                                                                                                                                                                                                        41
                                                                                                                                                                                                                                                                                                                                                                                                                                                         61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SerIleAsnGlyValLeuTrpThrValTyrHisGlyAlaGlyThrArgThrIleAlaSer
                                                                                                                                                                                                                                                                                                                      GlySerValValIleValGlyArgIleAsnLeuSerGlyAsp-----ThrAlaTyr
                                                                                                                                                                                                                                                                                                                                                                                      22 AlaGlnGlnThrArgGlyGluGlnGlyCysGlnLysThrSerHisThrGlyArgAspLys
                                                                                                                                                                                                                                                                                                                                                                                                                                                       42 AsnGlnValGluGlyGluValGlnIleValSerThrAlaThrGlnThrPheLeuAlaThr
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               (pases 1 to 1998)
Malcolm, B.A., Taremi, S.Shane., Weber, P.C. and Yao, N.
Single-chain recombinant complexes of hepatitis C virus
protease and NS4A cofactor peptide
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Single-chain recombinant complexes of hepatitis C protease and NS4A cofactor peptide
Patent: US 6211318-A 108 03 APR-2001,
Location/Qualifiers
1. 1998
                                                                                                                                                                  1998
167
11
11
                                                                                                                                                                                    Matches:
Conservative:
Mismatches:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1998 bp C
Sequence 103 from patent US 6211338.
AR145262
                                                                                                     424
                                                                                                                                                                                                                                       Indels:
                                                                                                                                                                    Length:
                                                                                                                                                                                                                                                      Gaps:
                                                                                                                                                                                                                                                                                      US-09-965-594-20 (1-197) x AR145267 (1-1998)
                                                                                  /organism="unknown"
596 c 568 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              AR145262.1 GI:15107129
                                                                                                                                                                 2.73e-65
878.50
92.86%
85.20%
```

us-09-965-594-20.rge

```
Unclassified
                                                                        Percent Similarity:
Best Local Similarity:
                  120 a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Unknown
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Unknown
                                               Alignment Scores:
                                                                                                                                                                                                                                                                                     82
                                                                                           Query Match:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             source
                 BASE COUNT
ORIGIN
                                                       ..
₽
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     DEFINITION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ACCESSION
VERSION
KEYWORDS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ORGANISM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             AUTHORS
TITLE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     REFERENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         JOURNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           RESULT 9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  AR145253
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  PEATURES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SOURCE
                                                                 Score
                                                       Pred
                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                            a
                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                 a
                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                      ò
                                                                                                                                                       ద
                                                                                                                                                                         ŏ
                                                                                                                                                                                         a
                                                                                                                                                                                                             ò
                                                                                                                                                                                                                              q
                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                     ö
                                                                                                                                                                                                                                                                                                    a
                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               PAT 08-AUG-2001
                                                                                                                                                                                                                                                                                                                                                                                       543
                                                                                                                                                                                                                                                                                                                                                                                                                                                  LeuSerProArgProlleSerTyrLeuLysGlySerSerGlyGlyProLeuLeuCysPro 161
                                                                                                                                                                 243
                                                                                                                                                                                                                                                                                                    ProLysGlyProValThrGlnMetTyrThrAsnValAspLysAspLeuValGlyTrpGln 101
                                                                                                                                                                                                                                                                                                                                                                             ValThrArgHisAlaAspValIleProValArgArgArgGlyAspSerArgGlySerLeu 141
                                                                                                                                                                                                                                                                                                                                          AlaProGlnGlySerArgSerLeuThrProCysThrCysGlySerSerAspLeuTyrLeu 121
                                                                                                                                                                                                                               61
                                                                                                                                                                                                                                                                    81
                                                                                                                                                                                            AlaGlnGlnThrArgGlyGluGlnGlyCysGlnLysThrSerHisThrGlyArgAspLys 41
                                                                                                                                                                                                                              SerIleAsnGlyValLeuTrpThrValTyrHisGlyAlaGlyThrArgThrIleAlaSer
                                                                                                                                                                                                                                                                                    244 TGCGTCAACGGCGTGTGGACCGTTTACCATGGTGCTGGCTCAAAGACCTTAGCCGGC
                                                                                                                                                                                                                                                                                                                                                                                                                         GlySerValValIleValGlyArglleAsnLeuSerGlyAsp-----ThrAlaTyr
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         NS3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1 (bases 1 to 651)
Malcolm,B.A., Taremi,S.Shane., Weber,P.C. and Yao,N.
Single-chain recombinant complexes of hepatitis C virus profease and NGA cofactor peptide
Patent: US 6211338-A 99 03-APR-2001;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      linear
                                                                       Length:
Matches:
Conservative:
Mismatches:
Indels:
Patent: US 6211338-A 103 03-APR-2001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             AR145258 651 bp
Sequence 99 from patent US 6211338.
AR145258
                                    424
                                                                                                                     Gaps:
                                                                                                                                      (1-1998)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Location/Qualifiers
         Location/Qualifiers
                 1. .1998
/organism="unknown"
596 c 568 q
                                                                                                                                      US-09-965-594-20 (1-197) x AR145262
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        AR145258.1 GI:15107125
                                                                       3.33e-65
877.50
92.86%
85.20%
86.03%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Unclassified.
                                                                                         Percent Similarity:
Best Local Similarity:
                                    Ø
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Unknown.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Unknown
                                                               Alignment Scores:
                                                                                                                                                                                                             124
                                                                                                                                                                                                                               42
                                                                                                                                                                                                                                                                   62
                                                                                                                                                                                                                                                                                                                        304
                                                                                                                                                                                                                                                                                                                                         102
                                                                                                                                                                                                                                                                                                                                                          364
                                                                                                                                                                                                                                                                                                                                                                             122
                                                                                                                                                                                                                                                                                                                                                                                                                 142
                                                                                                                                                                                                                                                                                                                                                                                                                                 184
                                                                                                                                                                                                                                                                                                                                                                                                                                                   162
                                                                                                                                                                                                                                                                                                                                                                                                                                                                      544
                                                                                                                                                                                                                                                                                                                                                                                               124
                                                                                                           Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          RESULT 8
AR145258
LOCUS
DEFINITION
ACCESSION
KEYWORDS
SOURCE
                    source
                                                                        Pred. No.:
                                   BASE COUNT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ORGANISM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     REFERENCE
AUTHORS
TITLE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        JOURNAL
FEATURES
          FEATURES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    a
                                                                                                                                                        ò
                                                                                                                                                                        g
                                                                                                                                                                                        ò
                                                                                                                                                                                                         g
                                                                                                                                                                                                                              õ
                                                                                                                                                                                                                                           a
                                                                                                                                                                                                                                                                  50
                                                                                                                                                                                                                                                                                q
                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                  a
                                                                                                                                                                                                                                                                                                                                         à
                                                                                                                                                                                                                                                                                                                                                      a
                                                                                                                                                                                                                                                                                                                                                                                            a
                                                                                                                                                                                                                                                                                                                                                                                                                               q
                                                                                                                                                                                                                                                                                                                                                                                                                                                   õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                    qd
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                 ò
```

```
08-AUG-2001
                                                                                                                                                                                                                                                                 141
                                                                                                                                                                                                                                                                                                                                                                                                            483
                                                                                                                                                                                                                                                                                                                                                                                                                                                                     161
                                                                                                                                                      ProLysGlyProValThrGlnMetTyrThrAsnValAspLysAspLeuValGlyTrpGln 101
                                                                                                                                                                                                                           61
                                                                                                                                                                                   22 AlaGlnGlnThrArgGlyGluGlnGlyCysGlnLysThrSerHisThrGlyArgAspLys 41
                                                                                                                                                                                                                         GlySerValValIleValGlyArglleAsnLeuSerGlyAsp-----ThrAlaTyr
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1 (bases 1 to 651)
Malcolm,B.A., Tarem1,S.Shane., Weber,P.C. and Yao,N.
Single-chain recombinant complexes of hepatitis C virus protease and NS4A cofactor peptide
Patent: US 6211338-A 94 03-APR-2001;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           PAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   648
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             182 ValAspPheIleProValGluSerLeuGluThrThrMetArgSer 196
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           linear
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        651
166
10
10
                                                  Length:
Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ONA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         AR145253 651 bp 1
Sequence 94 from patent US 6211338.
AR142253 GI:15107120
           144
                                                                                                                        (1-651)
/organism-"unknown"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                /organism-"unknown'
            σ
          200
                                                                                                                        US-09-965-594-20 (1-197) x AR145258
                                               1.35e-65
875.50
93.33%
85.13%
85.83%
         187 c
```

```
PAT 08-AUG-2001
                                                                                                                                                                                                                                                                                                                                                            161
                                                                                                                                                                                                                                                                                                                                                                                                                                                   AlaGlyHisAlaValGlyIlePheArgAlaAlaValSerThrArgGlyValAlaLySAla 181
                                                                                                                          304 CCAAAGGGGCCAATCACCCAGATGTACACTAATGTGGACCAGGACCTCGTCGGCTGGCAG 363
                                                                                                                                                                                                                                                                                                                        AlaProGlnGlySerArgSerLeuThrProCysThrCysGlySerSerAspLeuTyrLeu 121
                                                                                                                                                                           41
                                                                                                                                                                                                        42 AsnGlnValGluGlyGluValGlnIleValSerThrAlaThrGlnThrPheLeuAlaThr 61
                                                                                                                                                                                                                                             62 SerIleAsnGlyValLeuTrpThrValTyrHisGlyAlaGlyThrArgThrIleAlaSer 81
                                                                                                                                                                                                                                                                                                                                                                                                  22 AlaGlnGlnThrArgGlyGluGlnGlyCysGlnLysThrSerHisThrGlyArgAspLys
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1 (bases 1 to 1998)
Malcolm,B.A., Taremi,S.Shane., Weber,P.C. and Yao,N.
Single-chain recombinant complexes of hepatitis C virus NS3
protease and NS4 cofactor peptide
Patent: US 6211338-A 107 03-APR-2001;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                182 ValAspPheIleProValGluSerLeuGluThrThrMetArgSer 196
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         linear
                                       651
167
14
11
3
                                      Length:
Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Sequence 107 from patent US 6211338. ARI45266 ARI45266.1 GI:15107133
 145
                                                                                                         (1-651)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     /organism="unknown"
596 c 568 q
  σ
                                                                                                         US-09-965-594-20 (1-197) x AR145253
                                     1.64e-65
874.50
92.82%
85.64%
85.74%
 υ
 188
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Unclassified.
                                                          Percent Similarity:
Best Local Similarity:
  ø
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Unknown.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Unknown
                               Alignment Scores:
Pred. No.:
                                                                                                                                                                                                                                                                                                                        102
                                                                                                                                                                                                                                                                                                                                                                                                                       484
                                                                                                                                                                                                                                                                                                                                                                                                    142
                                                                                                                                                                                                                                                                                   82
                                                                             Query Match:
DB:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                BASE COUNT
ORIGIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    REFERENCE
AUTHORS
TITLE
BASE COUNT
ORIGIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     RESULT 10
AR145266
LOCUS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           JOURNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      FEATURES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  q
                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                        셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                            ò
                                                                                                                                              đ
                                                                                                                                                                                    셤
                                                                                                                                                                                                                        qq
                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                   õ
                                                                                                                                                                   ò
                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                              õ
```

```
08-AUG-2001
                                                                                                                                                                                                                                                                                                                                                       ProLysGlyProValThrGlnMetTyrThrAsnValAspLySAspLeuValGlyTrpGln 101
                                                                                                                                                                                                                                                                                  161
                                                                                                                                                                                                                                                                                                                                                                                                                     543
                                                                                                                                                                                                                                                                                                                                                                                                                                         162 AlaGlyHisAlaValGlyIlePheArgAlaAlaValSerThrArgGlyValAlaLysAla 181
                                                                                                                                                        64 GGITCTGTIGTITITITITIGITGGAATIATITIAICTGGTAGTGGTAGTAIAGCAACTAC 123
                                                                                                                                             41
                                                                                                                                                                                       61
                                                                                                                                                                                                                                                                                                                                                                                                 LeuSerProArgProlleSerTyrLeuLysGlySerSerGlyGlyProLeuLeuCysPro
                                                                                                                                                                                                                                                                                                                                                                                                           22 AlaginginThrArgGlyGluGlnGlyCysGlnLysThrSerHisThrGlyArgAspLys
                                                                                                                                                                                      42 AsnGlnValGluGlyGluValGlnIleValSerThrAlaThrGlnThrPheLeuAlaThr
                                                                                                     5 GlySerValValIleValGlyArglleAsnLeuSerGlyAsp-----ThrAlaTyr
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Unclassified.

1 (bases 1 to 1998)

1 (bases 1 to 1998)

Malcolm, B.A., Taremi, S.Shane., Weber, P.C. and Yao, N. Single-chain recombinant complexes of hepatitis C virus NS3 protease and NSAA cofactor peptide
Patent: US 6211338-A 102 03-APR-2001;
Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              PAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                               linear
        1998
166
16
11
         Length:
Matches:
Conservative:
Mismatches:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              DNA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            AR145261 1998 bp C Sequence 102 from patent US 6211338. AR145261.1 GI:15107128
                                                 Indels:
                                                                                US-09-965-594-20 (1-197) x AR145266 (1-1998)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     /organism-"unknown'
597 c 567 a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   6
         5.98e-65
874.50
92.86%
84.69%
85.74%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            .1998
                             Percent Similarity:
Best Local Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ø
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Unknown.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Unknown
Alignment Scores:
                                                                                                                                                                                                                                                                       83
                                                                                                                                                                                                                                                                                                                                                         122
                                                                                                                                                                                                                                                                                                                                                                            424
                                                                                                                                                                                                                                                                                                                                                                                                                     484
                                                                                                                                                                                                                                                                                                                                                                                                 142
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               source
                                                    Query Match:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       DEFINITION
ACCESSION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  BASE COUNT
ORIGIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ORGANISM
         Pred. No.:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          AUTHORS
TITLE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             VERSION
KEYWORDS
SOURCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                REFERENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         JOURNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          RESULT 11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    AR145261
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    FEATURES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                LOCUS
                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                             ò
                                                                                                                                                                 g
                                                                                                                                                                                      ŏ
                                                                                                                                                                                                       g
                                                                                                                                                                                                                               õ
                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                       õ
                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                               \delta
                                                                                                                                                                                                                                                                                                                               q
                                                                                                                                                                                                                                                                                                                                                        à
                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                 δ
                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                     ò
```

Alignment Scores:

Percent Similarity: 92.35% Conservative: 14 Best Local Similarity: 85.20% Mismatches: 12 Query Match: 85.64% Indels: 3 DB: 6 Gaps: 1 US-09-965-594-20 (1-197) x AR145269 (1-2016)		502 502 502 182 182 622 7 AF 7 PP 7 PP 1 DIG 103 103 103 103 103 103 103 103 103 103	PUBMED 10864644  REFERENCE 2 (bases 1 to 12734)  AUTHORS Lai, V.C.H. and Hong, Z.  TITLE Direct Submission  JOURNAL Submitted (16-MAY-2000) Antiviral Therapy, Schering-Plough Research Institute, 2015 Galloping Hill Road, Kenilworth, NJ 07033-0539, USA  FEATURES  Location/Qualifiers  source  112734
Pred. No.:         7.27e-65         Length:         1998           Score:         873.50         Matches:         167           Percent Similarity:         92.35%         Conservative:         14           Best Local Similarity:         85.20%         Mismatches:         12           Query Match:         6         Gaps:         1	5 GlySerValValileValGlyArgileAsnLeuSerGlyAspThralaTyr	42 LeuSerProA 42 CTCTCCCCA 62 AlaGlyHish 63 AlaGlyHish 64 TCGGGGCACG 82 ValAspPhe I 64 GTGGCATTG 64 GTGGACTTTG 65 AlaSS 66 ARI45269 67 GTGGACTTTG 67 GTGGACTTTG 68 ARI45269 68 ARI45269 69 ARI45269 69 ARI45269 60	source       12016         DASE COUNT       401 cm (12 am 603 cm 570 gm 431 cm 603 cm 570 gm 431 cm 603 cm 570 gm 431 cm 603 cm

5'UTR CDS

```
ENETGY TAUDISTS CHEGOVALUY OF CHILLY CONTINCE TO THOUGH THE PER ISS
EGYPERTACTENY TRILLENGY ELVY OF CHILLY OF AMOUNT AMOUNT FREILS
EGYPERTACTENY TRILLENGY ELVY OF THE CENTRAL CHILLY THE C
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       TEGIORAMICORGANISLIGIUMPERICTCOVESHIATDIELETTHGMOASEKTNYTCCR
TEGIORAMICORGANISLIGIUMPERICTCOVESHIATDIELETTHGMOASEKTNYTCCR
LOCHEMIKHGWCWMYNIEPWILVMIRTOCANLTEGOPPRECATTCRYORASDIAVYTOA
ROSPTPLTGCKKGKNFSFAGILARGPCNFEIAASDVLFKEHERISMEODTLYLVDGL
TNSLEGARGGTAKLTTHKKOVGILALGKKLENKSKTWFGAYAASPYCDVDDRIGYITAYDD
TNSLEGARGGTAKLTTHKKOVGILALGTRANTENKSTWFGRAYAASPYCDVDDRIGYITAYD
MYLLHES IPGSHOVDWCDCNTQLILYPELTTAADIESVLLLSLVVLSDFAPETASY
MYLLHES IPGSHOVDWCDCNTQLITAYBETTAALICLCLVTVRGWYOGILALLIT
GVOGHLOCK PEFSYALANDERIOGOGARGITTTAATITAFITANTAANTAADERIOGOGRAKIA
YLGRCTRETRY LAILHFRALPDISTSCHRESTSFNDTLATVVRTYRRSKFFPHRGGCIT
OKNIGEDDHNCILGGAWWTOPPODLYKGGSIESCKWCGYORKESPERPHYDIGKCKL
ENETGYRLVDSTSCHREGVAIVPOGTLKCKIGKTTVOVIAMDTKLGPHYPURSS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            GDVEKINGAÏSDYAAGGLEFVKSQAEKIKTAPLFKENAEAARGYVÖKFIDSLIENKEE
INTSCHWETTALKSIAARGHETAFALVLUKAAFGGESVSUHVKDAAVDLVYYY
HNKPSFPGDSETQORGERRYASLFISALAYTYTKHWYHNLSKVVEBALAYLPYATSA
LKMFTPTRLESVVILSTTYKTYLSIRKGKSDGLLGTGISAAMEILSQNPVSVGISVM
LGOGALAAMHAIESSEQKRTLJMKYWTKDLOAATDELVKSRPEKIIMALFEAVOTI
GNPLRLIYHLTGVYKKGWERKELSSRTAGRHLFTLIMFKRAPEKIIMALFRAVOTI
YILDLIYGLHKQINRGLKKWVLGWAPPFSCDWTPSDERIRLPTDNYLRVETRCPGGY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         AYTEVNEDVGTIKPVLGERVIPDPVVDINLQPEVQVDTSEVGITILGRETLMTTGVTP
VLEKVEPDASDNONSVKIGLDEGNYPGPGIQTHTLTEEIHNRDARPPIMILGSRNSIS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ALSLGQPKPKQVTKEAVRNLIEQKKDVEIPNWFASDDPVFLEVALKNDKYYLVGDVGE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             VKDQAKALGATDQTRI IKEVGSRTYAMKLSSWFLQASNKQMSLTPLFEELLLRCPPAT
KSNKGHMASAYQLAQGNWEPLGCGVHLGTIPARRVKIHPYEAYLKLKDFIEEEEKKPR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    AGIRLEKLPIVRAQTDTKTFHEAIRDKIDKSENRONPELHNKLLEIFHTIAOPTLKHT
YGEVTWEQLEAGINRKGAAGFLEKKNIGEVLDSEKHLVEQLYRDLKAGRKIKYYETAI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    EMKAFKNYGGKLTKVEESGPFLCRNRPGRGPVNYRVTKYYDDNLREIKPVAKLEGQVE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              HYYRGVTAKIDYSKGKMLLATDKWEVEHGVITRLAKRYTGVGFNGAYLGDEPNHRALV
ERDCATITKNTVQFLKMKKGCAFTYDLIISNLTRLIELVHRNNLEEKEIPTATVTTWL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   NRAKTARNINLYTGNDPREIRDLMAAGRMLVVALRDVDPELSEMVDFKGTFLDREALE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       VKDTVIREHNKWILKKIRFQGNLNTKKMLNPGKLSEQLDREGRKRNIYNHQIGTIMSS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     PKNEKRDVSDDWQAGDLVVEKRPRVIQYPEAKTRLAITKVMYNWVKQQPVVIPGYEGK
TPLFNIFDKVRKEWDSFNEPVAVSFDTKAMDTQVTSKDLQLIGEIQKYYYKKEWHKFI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SYSDTKEEGATKKKTOKPDRLERGKMKIVPKESEKDSKTKPPDATIVVEGVKYQVRKK
GRTKSKNTQDGLYHNKNKPQESRKRLEKALLAWAIIAIVLFQVTMGENITQWNLQDNG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  /translation-"melmTheGSGSVVIVGRIVLSGSGSITaCAQQTRGLLGCKITSL
TGRDKNQVEGEVQIVSTATOTFLATCINGVCWTVYHGAGTRTIASPKGPVIQMYTNVD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      QDLVGWPAPQGSRSLTPCTCGSSDLYLVTRHANVIPVRRRGDSRGSLLSPRPISYLKG
SSGGPLLCPAGHAVGLFRAAVCTRGVAKAVDFIPVENLETTTRSGSGADTEDVVCCSM
                                                                                                                                                                                                                                                                                                              /product-"polyprotein"
/protein_id-"AAF82566.1"
/db_xref-"G1:9049957"
/mol_type="genomic RN;
/db_xref="taxon:11099
                                                                                                                                                                                                                                                          /codon_start-1
                                                                                                                                       .385
```

```
DIIDHMTEVPVITADGEVIIRNGQRGSGQPDTSAGNSMLNVLTMMTAFCESTGVPYK
SFNRYARIHVGGDDGFLITEKGLGLKFANKGMOILHEAGKPOKITEGERKMYAYRFED
FECSHTPPVPVMSDNTSSHMAGNDFAVILSKMATRLDSGERGTTAFERAVEFSFLL
MYSWNPLVRRICGLLVLSQQPSTDPSKRATYYKGDPIGAYKDVIGRNLSELKPTGFER
LANLNLSLSTLGIWTKHTSKRIIQDCVAIGKEEGNWLVNADRLISSKTCHLYIDDKGF
TLQGRHYEGGLLKTETNPVMGVGTERYKLGPIVNLLLRRLKILLMTAVGVSS*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         PAT 08-AUG-2001
                                                                                                                                                                                                                                                                                                                                     892
                                                                                                                                                                                                           592
                                                                                                                                                                                                                                                                                                                                                                                         772
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         161
                                                                                                                                                                                                                                                                                                                                                                              ProLysGlyProValThrGlnMetTyrThrAsnValAspLysAspLeuValGlyTrpGln 101
                                                                                                                                                                                                                                                                                                                                                                                                                       AlaProGlnGlySerArgSerLeuThrProCysThrCysGlySerSerAspLeuTyrLeu 121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  AlaGlyHisAlaValGlyIlePheArgAlaAlaValSerThrArgGlyValAlaLysAla 181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      GCGGGACACGCCGTGGGCCTATTCAGGGCCGCGGTGTGCACCCGTGGAGTGGCCAAGGCG 952
                                                                                                                                                                                                                                                      41
                                                                                                                                                                                                                                                                                              19
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        22 AlaGlnGlnThrArgGlyGluGlnGlyCysGlnLysThrSerHisThrGlyArgAspLys
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          (bases 1 to 651)
Malcolm,B.A., Taremi,S.Shane., Weber,P.C. and Yao,N.
Single-chain recombinant complexes of hepatitis C virus NS3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ValAspPhelleProValGluSerLeuGluThrThrMetArgSer 196
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       others
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          linear
                                                                                                                          Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ONA
                                                                        Ų
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Sequence 98 from patent US 6211338.
ARI45257
ARI45257.1 GI:15107124
                                                                        2802
                                                                                                                 Length:
                                                                                                                                                                                        (1-12734)
                                                                                                                                                                   Gaps:
                                                                         ந
                                                                        3293
                                                                                                                                                                                         US-09-965-594-20 (1-197) x AF268278
                                                               .09. .12734
2608 c 33
                                                                                                                7.45e-64
872.50
91.79%
89.23%
85.54%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Unclassified.
                                                                                                                               Percent Similarity:
Best Local Similarity:
Query Match:
                                                                         ದ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Unknown.
                                                                        4030
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Unknown
                                                                                                       Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  162
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           182
                                                                                                                                                                                                                                 413
                                                                                                                                                                                                                                                                                             42
                                                                                                                                                                                                                                                                                                                  533
                                                                                                                                                                                                                                                                                                                                                                                                                       102
                                                                                                                                                                                                                                                                                                                                                                                                                                                                 122
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    773
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         142
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              833
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      893
                                                                                                                                                                                                             S
                                                                                                                                                                                                                                                                         473
                                                                                                                                                                                                                                                                                                                                                                               82
                                                                                                                                                                                                                                                                                                                                                                                                   653
                                                                                                                                                                                                                                                                                                                                                                                                                                             713
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               953
                                                             3'UTR
BASE COUNT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   DEFINITION
                                                                                                                Pred. No.:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ACCESSION
VERSION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ORGANISM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       AUTHORS
TITLE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     RESULT 14
AR145257
                                                                                                       Alignment
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  KEYWORDS
SOURCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             REFERENCE
                                                                                  ORIGIN
                                                                                                                            Score:
                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                   a
                                                                                                                                                                                                                                                                                                                                                                                                                     ð
                                                                                                                                                                                                                                                                                                                                                                                                                                           셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                              5
                                                                                                                                                                                                                                                                          윱
                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                       ò
```

```
Scores
                                                                                                                                                                                                                                                                                                                                                                                                                                                       182
                                                                                                                                                                                                            42
                                                                                                                                                                                                                                                                                                                102
                                                                                                                                                                                                                                                                                                                                                                   424
                                                                                                                                                                                                                                                                                                                                                                                                    484
                                                                                                                                                                                                                                                                                              104
          source
                           BASE COUNT
ORIGIN
                                                             ..
NO.
                                                      Alignment
  FEATURES
                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                  q
                                                                                                                                                                                                                                                                                                                                                                                                   a
                                                                                                                                                                                                                                                                                                                                                                                                                     á
                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                       á
                                                                                                                                                                                                                                                                               õ
                                                                                                                                                                                                                                                                                             q
                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                         ö
                                                                                                                                                         g
                                                                                                                                                                           à
                                                                                                                                                                                       g
                                                                                                                                                                                                            ò
                                                                                                                                                                                                                           q
                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           08-AUG-2001
                                                                                                                                                                                                                                            243
                                                                                                                                                                                                                                                                               303
                                                                                                                                                                                                                                                                                               101
                                                                                                                                                                                                                                                                                                                363
                                                                                                                                                                                                                                                                                                                                 121
                                                                                                                                                                                                                                                                                                                                                423
                                                                                                                                                                                                                                                                                                                                                                   141
                                                                                                                                                                                                                                                                                                                                                                                    483
                                                                                                                                                                                                                                                                                                                                                                                                    161
                                                                                                                                                                                                                                                                                                                                                                                                                     543
                                                                                                                                                                                                                                                                                                                                                                                                                                    183
                                                                                                                                                                                           41
                                                                                                                                                                                                                             61
                                                                                                                                                                                                                                                              81
                                                                                                                                                                                                                                                                                                                               ValThrargfisalaAspValIleProValArgArgArgGlyAspSerArgGlySerLeu
                                                                                                                                                                                                                                                                                                                                                                                                   64 GGTTCTGTTGTTATTGTTGGTAGAATTATTTATCTGGTAGTAGTAGTATCACGGCCTAC
                                                                                                                                                                                                    244 TGCGTCAACGGCGTGTGTTGGACCGTTTACCATGGTGCTGGCTCAAAGACCTTAGCCGC
                                                                                                                                                                                                                                                                                                        -----ThrAlaTyr
                                                                                                                                                                                          ProLysGlyProValThrGlnMetTyrThrAsnValAspLysAspLeuValGlyTrpGln
                                                                                                                                                                                                                                                                                                                                                                                    424 GTCACGAGACATGCTGACGTCATTCCGGTGCGCCGGCGGGGGGGACAGTAGGGGGGACGTG
                                                                                                                                                                                                                                                               SerileAsnGlyValLeuTrpThrValTyrHisGlyAlaGlyThrArgThrIleAlaSer
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1 (bases 1 to 651)
Malcolm, B.A., Taremi, S.Shane., Weber, P.C. and Yao, N.
Single-chain recombinant complexes of hepatitis C virus NS3
protease and NS4s of Oxfactor peptide
Patent: US 6211338-A 93 03-APR-2001;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           PAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                        182 ValAspPheIleProValGluSerLeuGluThrThrMetArgSer 196
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        648
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           linear
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 GIGGACTITGTGCCCGTAGAGTCCATGGAAACTACTATGCGGTCT
                                                                             651
166
15
11
11
                                                                                                                                                          5 GlySerValValIleValGlyArgIleAsnLeuSerGlyAsp-
                                                                             Length:
Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           DNA
protease and NS4A cofactor peptide
Patent: US 6211338-A 98 03-APR-2001;
Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Sequence 93 from patent US 6211338.
AR145252
AR145252.1 GI:15107119
                                                                                                                        Gaps:
                                                                                                                                         US-09-965-594-20 (1-197) x AR145257 (1-651)
                                 /organism-"unknown"
188 c 199 q
                                                                            2.95e-65
871.50
92.82%
85.13%
85.44%
                           .651
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Unclassified.
                                                                                                        Similarity:
                                             ø
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Unknown.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Unknown
                                                                                              Percent Similarity:
Best Local Similari
                                                                      Alignment Scores:
                                                                                                                                                                                                                                                               62
                                                                                                                                                                                                                                                                                               82
                                                                                                                                                                                                                                                                                                                                                                   122
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ARI45252
LOCUS
DEFINITION
ACCESSION
VERSION
                                                                                                                Query Match:
DB:
                             source
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ORGANISM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               REFERENCE
AUTHORS
                                                                                Pred. No.:
                                             BASE COUNT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         RESULT 15
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            KEYWORDS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 JOURNAL
           JOURNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                TITLE
                                                      ORIGIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    q
                                                                                                                                                                                                                                                                                                                                              Q
                                                                                                                                                                                                                                                                                                                                                                                 q
                                                                                                                                                                                                                                                                                                                                                                                                                                      ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  à
                                                                                                                                                                           g
                                                                                                                                                                                          ò
                                                                                                                                                                                                            a
                                                                                                                                                                                                                             δŏ
                                                                                                                                                                                                                                          qq
                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                             ά
                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                q
                                                                                                                                                          ò
```

```
243
                                                                                                                                                                                                                                                                                                                                                                                                                                             121
                                                                                                                                                                                                                                                                                                                                                                                                                                                         141
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      483
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              161
                                                                                                                                                                                                                      64 GGTTCTGTTGTTGTTGGTAGAATTATTTTATCTGGTAGTGGTAGTATATCACGGCCTAC 123
                                                                                                                                                                                                                                                                                                                                             62 SerileAsnGlyValLeuTrpThrValTyrHisGlyAlaGlyThrArgThrIleAlaSer 81
                                                                                                                                                                                                                                                                                                19
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              142 LeuSerProArgProlleSerTyrLeuLysGlySerSerGlyGlyProLeuLeuCysPro
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             122 ValThrArgHisAlaAspValIleProValArgArgArgGlyAspSerArgGlySerLeu
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          AlabroglnGlySerArgSerLeuThrProCysThrCysGlySerSerAspLeuTyrLeu
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CTCTCCCCCAGGCCTGTCTCCTACTTGAAGGCTCTTCGGGTGGTCCACTGCTCTGGCCT
                                                                                                                                                                                               GlySerValValIleValGlyArglleAsnLeuSerGlyAsp-----ThrAlaTyr
                                                                                                                                                                                                                                                                                                AsnGlnValGluGlyGluValGlnIleValSerThrAlaThrGlnThrPheLeuAlaThr
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           651
166
15
11
                                                                                                           Conservative:
Mismatches:
Indels:
                                                                                                 Matches:
                                     145
                                                                                                                                                  Gaps:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Search completed: August 31, 2003, 00:46:32 Job time : 2569.57 secs
                                                                                                                                                                        US-09-965-594-20 (1-197) x AR145252 (1-651)
Location/Qualifiers
            1. .651
/organism-"unknown"
                                      σ
                                                                                 3.59e-65
870.50
92.828
85.138
                                    188 C
                                                                                                                       Best Local Similarity:
Query Match:
                                      ಗ
                                                                                                            Percent Similarity:
```

ö

```
Hepatitis; NS3 protease; viral replication; chronic liver disease;
liver failure; liver cancer; mutant; mutein; ds.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Hepatitis C virus NS4A-NS3 fusion protease coding sequence #6.
is derived by analysis of the total score distribution
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     AL IGNMENTS
                                           SUMMARIES
                                                                                                                                                                                                                                                                                                                                     AAX80355
AAX80359
                                                                                                                                                                                                                                                                                                                                                                                                  AAX80345
AAX80358
                                                                                                                                                                                                                                                                                                                                                                                                                                                             AAX80349
ABX15706
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                AAX80352
AAX80360
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   AAV59361
ABK87285
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         AAX23259
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 AAX23258
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ABK87286
                                                                                                                                                                                                                                                                                                                 ABA95615
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     AAX80348
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             AAX80343
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        AAX80365
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 AAX80347
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          AAX80362
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                AAX80346
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      AAZ36212
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         AAN90327
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        AAN90336
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            AAX80357
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     AAX80351
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            AAV59364
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Location/Qualifiers
                                                                                                        ΩI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ВР
                                                                                                                                                                                                                              AAA73333 standard; DNA; 594
                                                                                                        DB
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (first entry)
                                                             %
Query
Match Length 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   651
651
1998
2016
648
648
650
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1933
9646
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  9646
12980
12980
16622
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         5360
6905
7310
7310
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      651
1998
1998
2013
                                                                                                                                                                                                                                                                                                                                                                                                651
1998
1998
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Hepatitis C virus.
Synthetic.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         19-DEC-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       868
867.5
867.5
866.5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          8588.5
8588.5
8588.5
8544.5
854.4
854.7
854.7
854.7
854.7
                                                                                                                                                                                                                                                                                                                                8885.5
8881.5
8878.5
8778.5
8778.5
8774.5
8774.5
8774.5
8774.5
8774.5
8774.5
8774.5
8774.5
8776.5
8776.5
8776.5
8776.5
8776.5
8776.5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              863.5
863.5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             AAA73333;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   858.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          RESULT 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Key
                                                                                                                                                                                                                                                                                                                                   Result
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          **MODEL-frame+, parameters.**
**CORD_L/GERME+, PARAMETERS.**
*
                                                                                                                                         August 30, 2003, 19:13:57 ; Search time 182.939 Seconds (without alignments) 2906.924 Million cell updates/sec
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           N_Geneseq_19Jun03:*

| SIDSI/gcgdata/geneseq/geneseqn-embl/NA1980.DAT:*
| SIDSI/gcgdata/geneseq/geneseqn-embl/NA1981.DAT:*
| SIDSI/gcgdata/geneseq/geneseqn-embl/NA1981.DAT:*
| SIDSI/gcgdata/geneseq/geneseqn-embl/NA1981.DAT:*
| SIDSI/gcgdata/geneseq/geneseqn-embl/NA1981.DAT:*
| SIDSI/gcgdata/geneseqy/geneseqn-embl/NA1986.DAT:*
| SIDSI/gcgdata/geneseqy/geneseqn-embl/NA1986.DAT:*
| SIDSI/gcgdata/geneseqy/geneseqn-embl/NA1986.DAT:*
| SIDSI/gcgdata/geneseqy/geneseqn-embl/NA1990.DAT:*
| SIDSI/gcgdata/geneseqy/geneseqn-embl/NA20018.DAT:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed,
                                                                                                                                                                                                                        US-09-965-594-20
1020
1 MKKKGSVVIVGRINLSGDTA.....VAKAVDFIPVESLETTMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     /SIDS1/gcgdata/geneseq/genesegn-emb1/NA2003.DAT:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   5105512
                 GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 /SIDS1/gcgdata/geneseg/genesegn-emb
                                                                                                    nucleic search, using frame_plus_p2n model
                                                                                                                                                                                                                                                                                                                                                                                                                                                                              hits satisfying chosen parameters:
                                                                                                                                                                                                                                                                                                                                                                                                                                       2552756 seqs, 1349719017 residues
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries
                                                                                                                                                                                                                                                                                                                                0.5
7.0
7.0
                                                                                                                                                                                                                                                                                                                                , Xgapext
, Ygapext
, Fgapext
, Delext
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Minimum DB seq length: 0
Maximum DB seq length: 2000000000
                                                                                                                                                                                                                                                                                                                            Xgapop 10.0,
Ygapop 10.0,
Fgapop 6.0,
Delop 6.0,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        line parameters:
                                                                                                                                                                                                                                                                                                             BLOSUM62
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Total number of
                                                                                                                                                                                                                              Title:
Perfect score:
                                                                                                                                                                                                                                                                                                           Scoring table:
                                                                                                      OM protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Database :
                                                                                                                                                                                                                                                                                                                                                                                                                                       Searched:
                                                                                                                                                                                                                                                                        Sequence:
```

Command

HCV NS4A-NS3 COMPL HCV NS4A-NS3

Combined open read Hepatitis C virus Combined open read

Composite hepatiti

Nucleotide sequenc

CDNA encoding hepa

Hepatitis C virus

Hepatitis C virus Hepatitis C virus

Hepatitis C virus Hepatitis C virus Hepatitis C virus Hepatitis C virus Hepatitis C virus Hepatitis C virus Chimeric BVDV/HCV HCV NS4A-NS3 compl

Hepatitis C virus Hepatitis C virus

Description

~

셤 à g

a ò

ò

g

ò

```
The present sequence is the coding sequence for a mutated version of a fusion protein created using the Hepatitis C virus (HCV) NS3 and NS4A protease enzymes. These proteins are both essential for the replication of the virus, acting to cleave its replicative proteins from the polyprotein produced from the HCV genome. Inhibitors of the two proteins should be effective as antiviral treatments of HCV infection. This is useful as HCV can lead to chronic liver disease such as cirrhosis, liver failure and liver cancer. The present invention concerns a number of NS3 mutants and NS3-NS4A fusion proteins which can be used to identify inhibitors of this type, as well as enabling structural studies of the protease and protease inhibitor complexes. The protein produced from this sequence contains the alpha-helix0-7 variant.
                                                                                                                                                                                                                                                                                                                             Modified hepatitis C virus (HCV) NS3 protease comprising at least l substitution of a hydrophobic alpha-helix 0 amino acid to a hydrophilic amino acid, useful for screening inhibitors that may treat hepatitis C
                                /product= "NS4A-NS3 fusion protein #6"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Sequence 594 BP; 104 A; 191 C; 152 G; 147 T; 0 other;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            594
197
0
0
                                                                                                                                                                                                                                           Weinheimer S, Zhang Y, Goldfarb V;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Conservative:
Mismatches:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Matches:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Length:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Indels:
                                                                                                                                                                                                          (BRIM ) BRISTOL-MYERS SQUIBB CO
                                                                                                                                                                                                                                                                                                                                                                                                                 Claim 26; F1g 16; 66pp; English
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            7.17e-87
1020.00
100.00%
100.00%
100.00%
                                                                                                                                       06-JAN-2000; 2000WO-US00345
                                                                                                                                                                       99US-0115271
                     Q
                /*tag-
                                                                                                                                                                                                                                                                          WPI; 2000-465976/40.
P-PSDB; AAB15224.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Percent Similarity:
Best Local Similarity:
Query Match:
                                                                   WO200040707-A1
                                                                                                                                                                                                                                           Wittekind M,
                                                                                                                                                                       08-JAN-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Alignment Scores:
                                                                                                      13-JUL-2000
```

: 80		21	Gaps:	0	
16-60-SN	55-594	US-09-965-594-20 (1-197) x AAA73333 (1-594)	594)		
ΟŸ	1	1 MctLysLysLysGlySerValValIleValGlyArgIleAsnLeuSerGlyAspThrAla	levalGlyArgile		20
qq	7	1 ATGAAAAAAAAGGATCGGTGTTATCGTCGGCCGTATCAACCTGTCCGGTGACACCGCT 60	TCGTCGGCCGTATC	AACCIGICCGGIGACACCGCI	09
Qy	21	<u></u>	InGlyCysGlnLys		40
ପ୍ପ	61	1 TAGGCTCAGCAGACTCGAGGTGGGGGGGGGGGGGGGGGG	AGGGTTGCCAGAAG	ACCICCCACACCGGICGIGAC	120
ολ	41	-	lnilevalSerThr		09
qq	121	1 AAAAACCAGGTTGAAGGTGAAGGTTCAGATCGTTTCCACCGCTACCCAGACCTTCCTGGCT	AGATCGTTTCCACC	SCTACCCAGACCTTCCTGGCT	180
δŷ	61	H -	ChryalTyrHisGly		80
qq	181	1 ACCICCATCAACGGIGITCIGIGGACCGITTACCACGGIGCIGGIACCCGTACCATCGCI	ACCGITIACCACGGT	GCTGGTACCCGTACCATCGCT	240
δλ	81	٠, -	(etTyrThrAsnVal	AspLysAspLeuValGlyTrp	100
qa	241	1 CCCCGAAAGGTCCGGTTACCCAGATGTACACCAACGTTGAAAAAAAA	TGTACACCAACGTT		300
0y	101	1 GlnAlaProGlnGlySerArgSerLeuThrProCysThrCysGlySerSerAspLeuTyr	euThrProCysThr	CysGlySerSerAspLeuTyr	120

```
420
                                                                                                                              ProAlaGlyHisAlaValGlyIlePheArgAlaAlaValSerThrArgGlyValAlaLys 180
                                          LeuValThrArgHisAlaAspVallleProValArgArgArgGlyAspSerArgGlySer 140
                                                                                                                                                                                                                                                                 481 CCGGCTGGTCACGCTGTTGGTATCTTCCGTGCTGTTTCCACCCGTGGTGTTGCTAAA 540
301 CAGGCTCCGCAGGGTTCCCGTTCCCTGACCCCGTGCACCTGCGGTTCCTCCGACCTGTAC 360
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Modified hepatitis C virus (HCV) NS3 protease comprising at least 1 substitution of a hydrophobic alpha-helix 0 amino acid to a hydrophilic amino acid, useful for screening inhibitors that may treat hepatitis C
                                                                      Hepatitis; NS3 protease; viral replication; chronic liver disease;
liver failure; liver cancer; mutant; mutein; ds.
                                                                                                                                                                                                                                                                                                              181 AlaValAspPheIleProValGluSerLeuGluThrThrMetArgSerPro 197
                                                                                                                                                                                                                                                                                                                                     Hepatitis C virus NS4A-NS3 fusion protease coding sequence #7.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          /*tag= a
/product= "NS4A-NS3 fusion protein #7"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Goldfarb V;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Zhang Y,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Location/Qualiflers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    8
                                                                                                                                                                                                                                                                                                                                                                                                                                                 BP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  (BRIM ) BRISTOL-MYERS SQUIBB
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Wittekind M, Weinheimer S,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           06-JAN-2000; 2000WO-US00345.
                                                                                                                                                                                                                                                                                                                                                                                                                                               AAA73334 standard; DNA; 594
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        99US-0115271
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 19-DEC-2000 (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      WPI; 2000-465976/40.
P-PSDB; AAB15225.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    WO200040707-A1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        08-JAN-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 13-JUL-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Synthetic.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       AAA73334;
                                          121
                                                                                                                                                                                                                         161
                                                                                      361
                                                                                                                                                                                                                                                                                                                                                                                                                            AAA7333
                                                                                                                                                                                                                                                                                                                                                                                                       RESULT
```

Claim 26; Fig 17; 66pp; English.

The present sequence is the coding sequence for a mutated version of a fusion protein created using the Hepatitis C virus (HCV) NS3 and NS4A protease enizymes. These proteins are both wesetital for the replication of the virus, acting to cleave its replicative proteins from the polyprotein produced from the HCV genome. Inhibitors of the two proteins should be effective as antiviral treatments of HCV infection. This is useful as HCV can lead to chronic liver disease such as cirrhosis, liver failure and liver cancer. The present invention concerns a number of NS3 mutants and NS3-NS4A fusion proteins which can be used to identify inhibitors of this type, as well as enabling structural studies of the protease and protease:inhibitor complexes. The protein produced from this sequence contains the alpha-helix0-7 variant.

 $\sim$ 

```
420
                                                                                                                                                                                                                                                                                                                                          480
                                                                                                                                                                                                                                                                                                                                                       240
                                                                                                                                                                                                                                                                 180
                                                                                                                                                                                                                                                          GlnAlaProGlnGlySerArgSerLeuThrProCysThrCysGlySerSerAspLeuTyr 120
                                                                                                                                                                                                                                                                                          LeuValThrArgHisAlaAspValIleProValArgArgArgGlyAspSerArgGlySer 140
                                                                                                                                                                                                                                                                                                                          LeuleuSerProArgProlleSerTyrLeuLysGlySerSerGlyGlyProLeuLeuCys 160
                                                                                               50
                                                                                                             9
                                                                                                                                                             9
                                                                                                                                                                                           80
                                                                                                                             40
                                                                                                                                      LysAsnGlnValGluGlyGluValGlnIleValSerThrAlaThrGlnThrPheLeuAla
                                                                                                                                                                    ThrSerIleAsnGlyValLeuTrpThrValTyrHisGlyAlaGlyThrArgThrIleAla
                                                                                             MetLysLysLysGlySerValValIleValGlyArgIleAsnLeuSerGlyAspThrAla
                                                                                                                             TyralaginginThrargglygluginglyCysglnLysThrSerHisThrGlyArgAsp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     protease; viral replication; chronic liver disease;
liver cancer; mutant; mutein; ds.
                                                                                                                                                                                                                                                                                                                                                                                      Hepatitis C virus NS4A-NS3 fusion protease coding sequence #5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             /product- "NS4A-NS3 fusion protein #5"
 other;
                        594
0
0
0
0
0
                      Length:
Matches:
Conservative:
Mismatches:
Indels:
0
Ë
G; 146
                                                                               US-09-965-594-20 (1-197) x AAA73334 (1-594)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Location/Qualifiers
1..594
/*tag= a
151
BP; 105 A; 192 C;
                                                                                                                                                                                                                                                                                                                                                                                                                                       ВР
                     6.24e-86
1010.00
99.49%
99.49%
                                                                                                                                                                                                                                                                                                                                                                                                                                       594
                                                                                                                                                                                                                                                                                                                                                                                                                                                                       (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                       DNA;
                                                                                                                                                                                                                                                                                                                                                                                                                                       standard;
                                     Percent Similarity:
Best Local Similarity:
Query Match:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Hepatitis C virus.
Synthetic.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Hepatitis; NS3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              liver failure;
Sequence 594
                                                                                                                                                                                                                                                                                                                                                                                                                                                                      19-DEC-2000
                Alignment Scores
                                                                                                                             21
                                                                                                                                                            41
                                                                                                                                                                           121
                                                                                                                                                                                                                                                         101
                                                                                                                                                                                                                                                                         301
                                                                                                                                                                                                                                                                                                                         141
                                                                                                                                                                                                                                                                                                                                                                                       181
                                                                                                                                                                                                                                                                                         121
                                                                                                                                                                                                                                                                                                                                        421
                                                                                                                                                                                                                                                                                                                                                                                                                                       AAA73332
                                                                                                                                                                                           61
                                                                                                                                                                                                                                                                                                                                                                                                                                                       AAA73332
                        Pred. No.:
                                                                                                                                                                                                                                                                                                                                                                                                                      RESULT 3
                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                      THE SOOK WAS ENDED IN
                                                                                              ò
                                                                                                            a
                                                                                                                             ð
                                                                                                                                          셤
                                                                                                                                                        ò
                                                                                                                                                                          a
                                                                                                                                                                                                                                        Q
                                                                                                                                                                                                                                                      õ
                                                                                                                                                                                                                                                                      음
                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                        õ
                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                 õ
                                                                                                                                                                                           à
                                                                                                                                                                                                        a
                                                                                                                                                                                                                         ò
```

```
The present sequence is the coding sequence for a mutated version of a fusion protein created using the Hepatitis C virus (HCV) NS3 and NS4A protease enzymes. These proteins are both essential for the replication of the virus, acting to cleave its replicative proteins from the polyprotein produced from the HCV genome. Inhibitors of the two proteins should be effective as antiviral treatments of HCV infection. This is useful as HCV can lead to chronic liver disease such as cirrhosis, liver failure and NS3-NS4A fusion proteins which can be used to identify inhibitors of this type, as well as enabling structural studies of the protease and protease:inhibitor complexes. The protein produced from this sequence contains the alpha-helix0-1 variant.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               100
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     GlnAlaProGlnGlySerArgSerLeuThrProCysThrCysGlySerSerAspLeuTyr 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             LeuValThrArgHisAlaAspValIleProValArgArgArgGlyAspSerArgGlySer 140
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     240
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            300
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CAGGCTCCGCAGGGTTCCCGTTCCCTGACCCCGTGCACCTGCGGGTTCCTCCGACCTGTAC 360
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        80
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           40
                                                                                                                                                                                                                    Modified hepatitis C virus (HCV) NS3 protease comprising at least 1 substitution of a hydrophobic alpha-helix 0 amino acid to a hydrophilic amino acid, useful for screening inhibitors that may treat hepatitis C
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           TyralaGlnGlnThrArgGlyGluGlnGlyCysGlnLysThrSerHisThrGlyArgAsp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ThrSerIleAsnGlyValLeuTrpThrValTyrHisGlyAlaGlyThrArgThrIleAla
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     MetLysLysLysGlyServalvalIleValGlyArgIleAsnLeuSerGlyAspThrAla
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                LysAsnGlnValGluGlyGluValGlnIleValSerThrAlaThrGlnThrPheLeuAla
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Sequence 594 BP; 105 A; 189 C; 153 G; 147 T; 0 other;
                                                                                                                                                    Goldfarb V;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Length:
Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         US-09-965-594-20 (1-197) x AAA73332 (1-594)
                                                                                                                                                     χ,
                                                                                                                                                    Zhang
                                                                                                                           8
                                                                                                                                                                                                                                                                                          Claim 26; Fig 15; 66pp; English
                                                                                                                        (BRIM ) BRISTOL-MYERS SQUIBB
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1.84e-85
1005.00
99.49%
98.48%
98.53%
                                                                                                                                                    Wittekind M, Weinheimer S,
                                                                  06-JAN-2000; 2000WO-DS00345
                                                                                              990S-0115271
                                                                                                                                                                            WPI; 2000-465976/40.
P-PSDB; AAB15223.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Best Local Similarity:
             WO200040707-A1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Percent Similarity:
                                                                                              08-JAN-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        241
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           21
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                41
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               301
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             81
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     101
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Query Match:
DB:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ..
2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Score
ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
```

g

a ö

ð 용 à

```
4.72e-84
990.00
97.97%
96.95%
97.06%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Hepatitis C virus.
                                                   Percent Similarity:
Best Local Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    19-DEC-2000
Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Synthetic
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  AAA73330;
                                                                                     Query Match:
                   ..
8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Key
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               AAA73330
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              RESULT
                   Pred.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                        셤
                                                                                                                                                                                                                                                                                                                                                                                                                                    셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                 ò
                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                         à
                                                                                                                                                                                                                                                                                                                              ö
                                                                                                                                                                                                                                                                                                                                                                셤
                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 The present sequence is the coding sequence for a mutated version of a fusion protein created using the Hepatitis C virus (HCV) NS3 and NS4A protease enzymes. These proteins are both essential for the replication of the virus, acting to cleave its replicative proteins from the polyprotein produced from the HCV genome. Inhibitors of the two proteins should be effective as antiviral treatments of HCV infection. This is useful as HCV can lead to chronic liver disease such as cirrhosis, liver failure and liver cancer. The present invention concerns a number of NS3 mutants and NS3-NS4A fusion proteins which can be used to identify inhibitors of this type, as well as enabling structural studies of the protease and protease: Inhibitor complexes. The protein produced from this
                                                     160
                                                                                       480
                                                                                                                           180
                                                                                                                                                              540
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Modified hepatitis C virus (HCV) NS3 protease comprising at least 1 substitution of a hydrophobic alpha-helix 0 amino acid to a hydrophilic amino acid, useful for screening inhibitors that may treat hepatitis C
                                                   CTGGTTACCCGTCACGCTGACGTTATCCCGGTTCGTCGTCGTGGTGACTCCCGTGGTTCC
Hepatitis; NS3 protease; viral replication; chronic liver disease;
liver failure; liver cancer; mutant; mutein; ds.
                                                                                                                                                                                                Hepatitis C virus NS4A-NS3 fusion protease coding sequence #4.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  /product- "NS4A-NS3 fusion protein #4"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Goldfarb V;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Wittekind M. Weinheimer S. Zhang Y.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Location/Qualifiers
1..594
/*tag= a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Claim 26; Fig 14; 66pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (BRIM ) BRISTOL-MYERS SQUIBB CO
                                                                                                                                                                                                                                                                                                            BP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            06-JAN-2000; 2000WO-US00345
                                                                                                                                                                                                                                                                                                            AAA73331 standard; DNA; 594
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               99US-0115271
                                                                                                                                                                                                                                                                                                                                                                                   (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    WPI; 2000-465976/40.
P-PSDB; AAB15222.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Hepatitis C virus.
Synthetic.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      WO200040707-A1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               08-JAN-1999;
                                                                                                                                                                                                                                                                                                                                                                                   19-DEC-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          13-JUL-2000
                                                                                       421
                                                                                                                                                                                                                                                                                                                                               AAA73331;
                   361
                                                       141
                                                                                                                                                                                                                                                                        RESULT 4
```

594 BP; 105 A; 187 C; 155 G; 147 T; 0 other;

Sequence

sequence contains the alpha-helix0-1 variant.

```
160
                                                                                                                                                                                                                                                                                                                                                                          480
                                                                                                                                                                                                                                        140
                                                                                                                                                                                                                                                                                                                                     420
                                                                                                                                                                                                                                                                                                                                                                                              161 ProAlaGlyHisAlaValGlyIlePheArgAlaAlaValSerThrArgGlyValAlaLys 180
                                                                                                                                                                                                                                                                                                                                                                                                               540
                                                                                                                                                                                                                                                                                                360
                                                                                                                                     180
                                                                                                                                                                                                                                                                               101 GlnAlaProGlnGlySerArgSerLeuThrProCysThrCysGlySerSerAspLeuTyr 120
                                                                                                                          40
                                                                                                                                                                                                     80
                                                                                                                                                                                                                                                                                                                 1 MetLysLysClysClySerValValIleValGlyArgIleAsnLeuSerGlyAspThrAla
                                                                                                       1 ATGAAAAAAAGGATCCGTTGTTATCGTCGGCCGTATCAACCTGTCCGGTGACACCGCT
                                                                                                                          21 TyralaginginthrargglygluginglycysglnLysthrSerHisThrGlyArgasp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Hepatitis; NS3 protease; viral replication; chronic liver disease;
liver failure; liver cancer; mutant; mutein; ds.
                                                                                                                                                                                                                                                                                                                                                                                                                                   181 AlavalAspPheIleProValGluSerLeuGluThrThrMetArgSerPro 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                      591
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Hepatitis C virus NS4A-NS3 fusion protease coding sequence #3
                                                                                                                                                                                                                                                                                                                                                                                                                                              /*tag= a
/product= "NS4A-NS3 fusion protein #3"
 594
191
2
4
0
0
Length:
Matches:
Conservative:
Mismatches:
                                      Indels:
                                                Gaps:
                                                                  (1-594)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Location/Qualifiers
                                                                  US-09-965-594-20 (1-197) x AAA73331
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           BP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           AAA73330 standard; DNA; 594
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1..594
/*tag=
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               WO200040707-A1
```

```
161 ProAlaGlyHisAlaValGlyIlePheArgAlaAlaValSerThrArgGlyValAlaLyS 180
                                                                                                    181 AlavalAspPheIleProvalGluSerLeuGluThrThrMetArgSerPro 197
                                                                                                                                                                                                                                                                                                                                                                                                                      /*tag= a
/product= *NS4A-NS3 fusion protein #8*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Sequence 594 BP; 98 A; 189 C; 153 G; 154 T; 0 other;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Goldfarb
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Length:
Matches:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Weinheimer S, Zhang Y,
                                                                                                                                                                                                                                                                                                                                                                                       Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Disclosure; Fig 18; 66pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                (BRIM ) BRISTOL-MYERS SQUIBB CO.
                                                                                                                                                                                       BP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     6.41e-80
946.00
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       06-JAN-2000; 2000WO-US00345.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   99US-0115271.
                                                                                                                                                                                       AAA73335 standard; DNA; 594
                                                                                                                                                                                                                                                (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        WPI; 2000-465976/40.
P-PSDB; AAB15226.
                                                                                                                                                                                                                                                                                                                                           Hepatitis C virus.
Synthetic.
                                                                                                                                                                                                                                                                                                       Hepatitis; NS3
                                                                                                                                                                                                                                                                                                                     liver failure;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                WO200040707-A1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Wittekind M,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   08-JAN-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Alignment Scores:
Pred. No.:
Score:
                                                                                                                                                                                                                                             19-DEC-2000
                                                                                                                                                                                                                    AAA73335;
                                                                                                                                                          AAA73335
                                                                                                                                                                                       The present sequence is the coding sequence for a mutated version of a fusion protein created using the Hepatitis C virus (HCV) NS3 and NS4A protease enzymes. These proteins are both essential for the replication of the virus, acting to cleave its replicative proteins from the POlyprotein produced from the HCV genome. Inhibitors of the two proteins should be effective as antiviral treatments of HCV infection. This is useful as HCV can lead to chronic liver disease such as cirrhosis, liver failure and liver cancer. The present invention concerns a number of NS3 mutants and NS3-NS4A fusion proteins which can be used to identify inhibitors of this type, as well as enabling structural studies of the protease and protease:inhibitor complexes. The protein produced from this sequence contains the alpha-helix0-1 variant.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          140
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SerProLysGlyProValThrGlnMetTyrThrAsnValAspLysAspLeuValGlyTrp 100
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 GlnAlaProGlnGlySerArgSerLeuThrProCysThrCysGlySerSerAspLeuTyr 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               80
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   9
                                                                                                                                                                        Modified hepatitis C virus (HCV) NS3 protease comprising at least 1 substitution of a hydrophobic alpha-helix 0 amino acid to a hydrophilic amino acid, useful for screening Inhibitors that may treat hepatitis C
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         LysAsnGlnValGluGlyGluValGlnIleValSerThrAlaThrGlnThrPheLeuAla
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ThrSerIleAsnGlyValLeuTrpThrValTyrHisGlyAlaGlyThrArgThrIleAla
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ACCTGCATCAACGGTGTTTGCTGGACCGTTTACCACGGTGCTGGTACCCGTACCATCGCT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                Sequence 594 BP; 103 A; 186 C; 156 G; 149 T; 0 other;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         294
1188
7
0
                                                                                                     Goldfarb V;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Length:
Matches:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Gaps:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              US-09-965-594-20 (1-197) x AAA73330 (1-594)
                                                                                                     Wittekind M, Weinheimer S, Zhang Y,
                                                                                                                                                                                                                                               Claim 26; Fig 13; 66pp; English.
                                                                          (BRIM ) BRISTOL-MYERS SQUIBB CO
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1.87e-82
973.00
96.45%
95.43%
                  06-JAN-2000; 2000WO-US00345.
                                             990S-0115271
                                                                                                                                2000-465976/40.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Best Local Similarity:
                                                                                                                                              P-PSDB; AAB15221
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Percent Similarity:
                                             08-JAN-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       81
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 101
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     361
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Query Match
```

a ŏ g ò

ð

õ g ò å ð 셤 ò

```
protease and protease inhibitor complexes. The protein produced from this sequence contains the alpha-helixO wild-type sequence.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    The present sequence is the coding sequence for a mutated version of a fusion protein created using the Hepatitis C virus (HCV) NS3 and NS4A protease enzymes. These proteins are both essential for the replication of the virus, acting to cleave its replicative proteins from the polyprotein produced from the HCV genome. Inhibitors of the two proteins should be effective as antiviral treatments of HCV infection. This is useful as HCV can lead to chronic liver disease such as cirrhosis, liver failure and liver cancer. The present invention concerns a number of NS3 mutants and NS3-NS4 fusion proteins which can be used to identify inhibitors of this type, as well as enabling structural studies of the
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Modified hepatitis C virus (HCV) NS3 protease comprising at least 1 substitution of a hydrophobic alpha-helix 0 amino acid to a hydrophilic amino acid, useful for screening inhibitors that may treat hepatitis C
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     protease; viral replication; chronic liver disease; liver cancer; mutant; muteln; ds.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Hepatitis C virus NS4A-NS3 fusion protease coding sequence #8
```

```
540
                                                                                                                                                                     180
                                                                                                                                                                                                            240
                                                                                                                                                                                                                                                                    120
                                                                                                                                                                                                                                                                                         360
                                                                                                                                                                                                                                                                                                            140
                                                                                                                                                                                                                                                                                                                                420
                                                                                                                                                                                                                                                                                                                                                                                          ProAlaGlyHisAlaValGlyIlePheArgAlaAlaValSerThrArgGlyValAlaLys 180
                                                                                                                       61 TACGCTCAGCAGACTCGAGGTCTGCATCATCATCATCCTCCTCGTGACCGGTGGAC 120
                                                                                                                                                                                                                               100
                                                                                                                                                                                                                                                  300
                                                                                                                                                  9
                                                                                                                                                                                         80
                                                                                        09
                                                                                                            40
                                                                                                                                                              301 CCGGCTCCGCAGGGTTCCCGTTCCCTGACCCCGTGCACCTGCGGTTCCTCCGACCTGTAC
                                                                                                                                                                                                                                                                                                                       TyralaGlnGlnThrargGlyGluGlnGlyCysGlnLysThrSerHisThrGlyArgAsp
                                                                                                                                                                                                                              SerProLysGlyProValThrGlnMetTyrThrAsnValAspLysAspLeuValGlyTrp
                                                                                                                                                                                                                                                                     GlnAlaProGlnGlySerArgSerLeuThrProCysThrCysGlySerSerAspLeuTyr
                                                                                                                                                                                                                                                                                                            LeuValThrArgHisAlaAspValIleProValArgArgArgGlyAspSerArgGlySer
                                                                               LysAsnGlnValGluGlyGluValGlnIleValSerThrAlaThrGlnThrPheLeuAla
                                                                                                                                                                                          ThrSerIleAsnGlyValLeuTrpThrValTyrHisGlyAlaGlyThrArgThrIleAla
                                                                                                                                                                                                            ACCIGCATCAACGGIGTTIGCIGGACCGTTIACCACGGIGCIGGIACCGGIACCAICGCI
                                                                     MetLysiysLysGlySerValVallleValGlyArgIlcAsnLeuSerGlyAspThrAla
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Hepatitis; NS3 protease; viral replication: chronic liver disease;
liver failure; liver cancer; mutant; mutein; ds.
                                                                                                                                                                                                                                                                                                                                                                                                                                            AlaValAspPheIleProValGluScrLeuGluThrThrMetArgSerPro 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Hepatitis C virus NS4A-NS3 fusion protease coding sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         /*tag= a
/product= "NS4A-NS3 fusion protein #2"
 0000
         Mismatches:
                      Indels:
                               Gaps:
                                                  (1-594)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Location/Qualiflers
                                                  US-09-965-594-20 (1-197) x AAA73335
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           я
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           AAA73329 standard; DNA; 588
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             06-JAN-2000; 2000WO-US00345
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (first entry)
94.42%
94.42%
92.75%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ...588
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Hepatitis C virus.
Synthetic.
         Best Local Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      W0200040707-A1
 Percent Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  19-DEC-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          13-JUL-2000
                                                                                                            21
                                                                                                                                                                                                            181
                                                                                                                                                                                                                                                                                                                                361
                                                                                                                                                                                                                                                                                                                                                                      421
                                                                                                                                                                                                                                                                                                                                                                                         161
                                                                                                                                                                                                                                                                                                                                                                                                                                 181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               AAA73329;
                                                                                                                                                   41
                                                                                                                                                                                          61
                                                                                                                                                                                                                                                                    101
                                                                                                                                                                                                                                                                                                            121
                                                                                                                                                                                                                                                                                                                                                                                                             481
                                                                                                                                                                                                                                81
                  Query Match:
DB:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                         RESULT 7
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                            ò
                                                                                                                               g
                                                                                                                                                ò
                                                                                                                                                                    qq
                                                                                                                                                                                         ò
                                                                                                                                                                                                        셤
                                                                                                                                                                                                                                                                 ð
                                                                                                                                                                                                                                                                                      q
                                                                                                                                                                                                                                                                                                            ŏ
                                                                                                                                                                                                                                                                                                                            a
                                                                                                                                                                                                                                                                                                                                               ç
                                                                                                                                                                                                                                                                                                                                                                   q
                                                                                                                                                                                                                                                                                                                                                                                         á
                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                 ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                  a
                                                                      õ
                                                                                                                                                                                                                            à
                                                                                                                                                                                                                                                g
```

```
The present sequence is the coding sequence for a mutated version of a fusion protein created using the Hepatitis C virus (HCV) NS3 and NS4A protease enzymes. These proteins are both essential for the replication of the virus, acting to cleave its replicative proteins from the polyprotein produced from the HCV genome. Inhibitors of the two proteins should be effective as antiviral treatments of HCV infection. This is useful as HCV can lead to chronic liver disease such as cirrhosis, liver failure and liver cancer. The present invention concerns a number of NS3 mutants and NS3-NS4A fusion proteins which can be used to identify inhibitors of this type, as well as enabling structural studies of the protease and protease:inhibitor complexes. The protein produced from this sequence contains the alpha-helix0-1 variant.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         121 LeuValThrArgHisAlaAspValIleProValArgArgArgGlyAspSerArgGlySer 140
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           174
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            295 CCGGCTCCGCAGGGTTCCCGTTCCCTGACCCCGTGCACCTGCGGTTCCTCCGACCTGTAC 354
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   61 ThrSerileAsnGlyValLeuTrpThrValTyrHisGlyAlaGlyThrArgThrIleAla 80
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             54
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             40
                                                                                                                                                 Modified hepatitis C virus (HCV) NS3 protease comprising at least 1 substitution of a hydrophobic alpha-helix 0 amino acid to a hydrophilic amino acid, useful for screening inhibitors that may treat hepatitis C
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              21 TyralaginginthrargglygluginglyCysGlnLysThrSerHisThrGlyArgAsp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          41 LysAsnGlnValGluGlyGluValGlnIleValSerThrAlaThrGlnThrPheLeuAla
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        175 ACCIGCATCAACGGIGITIGCIGGACCGITIACCACGGIGCIGGIACCCGIACCAICGCI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             101 GlnAlaProGlnGlySerArgSerLeuThrProCysThrCysGlySerSerAspLeuTyr
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Sequence 588 BP; 103 A; 180 C; 156 G; 149 T; 0 other;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          588
1184
3
2
                                                                       >
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Length:
Matches:
Conservative:
Mismatches:
Indels:
                                                                         Goldfarb
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Gaps:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  US-09-965-594-20 (1-197) x AAA73329 (1-588)
                                                                         Zhang Y,
                                           8
                                                                                                                                                                                                                               Claim 26; Fig 12; 66pp; English.
                                          (BRIM ) BRISTOL-MYERS SQUIBB
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        2.88e-79
939.00
94.92%
93.40%
                                                                         Weinheimer S,
           99US-0115271
                                                                                                      2000-465976/40.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Percent Similarity:
Best Local Similarity:
                                                                                                                     P-PSDB; AAB15220
                                                                         Wittekind M,
             08-JAN-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Alignment Scores:
Pred. No.:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Query Match:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ..
Q
ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
```

~

RESULT g

셤

```
ProAlaGlyHisAlaValGlyIlePheArgAlaAlaValSerThrArgGlyValAlaLys 180
                                                                                                                                                                                                                                                   SerProLysGlyProValThrGlnMetTyrThrAsnValAspLysAspLeuValGlyTrp 100
                                                                                                                                                                                                                                                                  MetLysLysLysClySerValValileValGlyArgileAsnLeuSerGlyAspThrAla
                                                                                                                                                                                           LeuLeuSerProArgProIleSerTyrLeuLysGlySerSerGlyGlyProLeuLeuCys
                                                                                                                                                        1 ATGAAAAAAAAGGTTCCGTTGTTATCGTCGCCGTATAGTACTGAACGGT-----GCT
                                                                                  21 TyrAlaGlnGlnThrArgGlyGluGlnGlyCysGlnLysThrSerHisThrGlyArgAsp
                                                                                                                                          41 LysasnGlnValGluGlyGluValGlnIleValSerThrAlaThrGlnThrPheLeuAla
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Pestivirus; Npro; protease; NS3; screening; ds.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Chimeric - Bovine viral diarrhea virus Chimeric - Repatitis C virus.
 US-09-965-594-20 (1-197) x AAA73328 (1-588)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Chimeric BVDV/HCV NS3-wt sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                BP.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ABA95615 standard; DNA; 12734
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    99US-0344456
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Lau JYN;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (SCHE ) SCHERING CORP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               WPI; 2002-121103/16
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Lai VCH,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  25-JUN-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    21-MAR-2002
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             US6326137-B1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            25-JUN-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        04-DEC-2001
                                                                                                                                                                                                                                                                                                                                                                                                                   141
                                                                                                                                                                                                                                                                                                                                                                                                                                             415
                                                                                                                                                                                                                                                                                                                                                                                                                                                                         161
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 475
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        535
                                                                                                                                                                                                                                                   81
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ABA95615
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Hong 2,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   RESULT 9
ABA95615
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ö
                                                     셤
                                                                                ò
                                                                                                          đ
                                                                                                                                       ò
                                                                                                                                                                g
                                                                                                                                                                                           ò
                                                                                                                                                                                                                     a
                                                                                                                                                                                                                                                  õ
                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                        à
                                                                                                                                                                                                                                                                                                                                 q
                                                                                                                                                                                                                                                                                                                                                            ö
                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                         Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ô
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        셤
the present invention concerns a number of NS3 mutants and NS3-NS4A fusion proteins which can be used to identify inhibitors of this type, as
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       The present sequence is the coding sequence for a fusion protein created using the Hepatitis C virus (HCV) NS3 and NS4A protease enzymes. These proteins are both essential for the replication of the virus, acting to cleave its replicative proteins from the polyprotein produced from the HCV genome. Inhibitors of the two proteins should be effective as antiviral treatments of HCV infection. This is useful as HCV can lead to chronic liver disease such as cirrhosis, liver failure and liver cancer.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Modified hepatitis C virus (HCV) NS3 protease comprising at least 1 substitution of a hydrophobic alpha-helix 0 amino acid to a hydrophilic amino acid, useful for screening inhibitors that may treat hepatitis C
                                                                                                                                                                                                                                                Hepatitis: NS3 protease; viral replication; chronic liver disease;
liver failure; liver cancer; ds.
                                                     Hepatitis C virus NS4A-NS3 fusion protease coding sequence #1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   well as enabling structural studies of the protease and
protease:inhibitor complexes.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      588
1182
122
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Sequence 588 BP; 97 A; 183 C; 153 G; 155 T; 0 other;
                                                                                                                                                                                                                                                                                                                                                              /*tag= a
/product= "NS3-NS4A fusion protein"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Wittekind M, Weinheimer S, Zhang Y, Goldfarb V;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Conservative:
Mismatches:
Indels:
Gaps:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Length:
Matches:
                                                                                                                                                                                                                                                                                                                                 Location/Qualifiers
1..588
/*tag= a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Disclosure; Fig 10; 66pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               (BRIM ) BRISTOL-MYERS SQUIBB CO
                                                                                                                                       ВР
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  9.89e-77
912.00
92.89%
92.39%
89.41%
                                                                                                                                       588
                                                                                                                                                                                                                                                                                                                                                                                                                                                         06-JAN-2000; 2000WO-US00345
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    99US-0115271
                                                                                                                                                                                              (first entry)
                                                                                                                                       DNA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   WPI; 2000-465976/40.
P-PSDB; AAB15212.
                                                                                                                                       standard;
                                                                                                                                                                                                                                                                                         Hepatitis C virus. Synthetic.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Percent Similarity:
Best Local Similarity:
Query Match:
DB:
                                                                                                                                                                                                                                                                                                                                                                                                   WO200040707-A1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    08-JAN-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Alignment Scores:
                                                                                                                                                                                            19-DEC-2000
                                                                                                                                                                                                                                                                                                                                                                                                                               13-JUL-2000
                         475
                                                     181
161
                                                                                                                                       AAA73328
                                                                                                                                                                 AAA73328
```

160

234

174 80

9

40

474

534

Nucleic acid construct encoding chimeric Hepatitis C Virus (HCV)

œ

```
The present invention relates to a nucleic acid construct encoding a chimeric Hepatitis C virus (HCV)-pestivirus genome. The construct comprises a pestivirus genome where a Npro pestivirus protease gene is replaced with a genome encoding a functional HCV NS3 protease. Furthermore, each junction site recognised by the Npro protease is replaced with a junction site recognised by the HCV NS3 protease. The construct is useful for screening compounds that inhibit HCV in vivo by inhibiting HCV protease, where screening may be in cell culture or in an animal model. The present sequence is a chimeric clone of BVDV (bovine viral diarrhea virus)/HCV NS3-wt, which was used to illustrate the present invention.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    592
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       772
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  113 GGTAGTGTTGTTGTTGTTGGTAGAATTGTTTTATCTGGTAGTGGTAGTATCACGGCGTAC 472
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              593 TGCATCAATGGGGTATGCTGGACTGTCTACCACGGGGCCGGAACGAGGACCATCGCATCA 652
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               82 ProLysGlyProValThrGlnMetTyrThrAsnValAspLysAspLeuValGlyTrpGln 101
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  102 AlaProGlnGlySerArgSerLeuThrProCysThrCysGlySerSerAspLeuTyrLeu 121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             LeuSerProArgProlleSerTyrLeuLysGlySerSerGlyGlyProLeuLeuCysPro 161
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              162 AlaGlyHisAlaValGlyIlePheArgAlaAlaValSerThrArgGlyValAlaLysAla 181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   893 GCGGGACACGCCGTGGGCCTATTCAGGGCCGCGGTGTGCACCCGTGGAGTGGCCAAGGC6 952
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       47
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            81
 pestivirus genome where the Npro protease gene is replaced with NS3 protease gene, useful for in vivo screening of compounds which inhibit
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           42 AsnGlnValGluGlyGluValGlnIleValSerThrAlaThrGlnThrPheLeuAlaThr
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SerIleAsnGlyValLeuTrpThrValTyrKisGlyAlaGlyThrArgThrIleAlaSer
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                653 CCCAAGGGTCCTGTCATCCAGATGTATACCAATGTGGACCAAGACCTTGTGGGCTGGCCC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      5 GlySerValvalileValGlyArgileAsnLeuSerGlyAsp------ThrAlaTyr
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   22 AlaGlnGlnThrArgGlyGluGlnGlyCysGlnLysThrSerHisThrGlyArgAspLys
                                                                                                                                                                                                                                                                                                    Sequence 12734 BP; 4032 A; 2604 C; 3295 G; 2803 T; 0 other;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      182 ValAspPheIleProValGluSerLeuGluThrThrMctArgScr 196
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Length:
Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                               US-09-965-594-20 (1-197) x ABA95615 (1-12734)
                                                                      Example 2; Columns 17-28; 20pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                               Gaps:
pestivirus genome where the Npro
                                                                                                                                                                                                                                                                                                                                                   3.12e-73
892.50
92.82%
90.77%
87.50%
                                                                                                                                                                                                                                                                                                                                                                                                              Similarity:
                                                                                                                                                                                                                                                                                                                                                                                            Percent Similarity:
                                      HCV infection
                                                                                                                                                                                                                                                                                                                                      Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             142
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            62
                                                                                                                                                                                                                                                                                                                                                                                                                             Query Match:
DB:
                                                                                                                                                                                                                                                                                                                                                          Pred. No.:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          RESULT 10
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       AAX80355
ID AAX8
XX
 ద
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            윤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   윤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           DD
```

```
The present invention describes a covalent hepatitis C virus (HCV) NS4A-NS3 complex comprising a central hydrophobic domain of native HCV NS4A-NS3 complex comprising a central hydrophobic domain. Where the hydrophobic domain of native HCV NS3 perine protease domain, where the hydrophobic domain of native HCV NS4A peptide is tethered by the linker to the amino terminus of the HCV NS3A protease domain. The present Sequence encodes an example of the above complex. The covalent NS4A-NS3 complexes are useful for structural determination and determination of mode of binding of HCV inhibitors by NMR spectroscopy. They can also be used for detecting inhibitors of the protease activity, the helicase activity and the ATPASS activity of NS3. The covalent covalent covalent covalent protease-peptide complexes previously available.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   GlySerValValIleValGlyArgIleAsnLeuSerGlyAsp-----ThrAlaTyr
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                22 AlaGlnGlnThrArgGlyGluGlnGlyCysGlnLysThrSerHisThrGlyArgAspLys
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              124 TCCCAACAGACGCGGGGCCTACTTGGTTGCAAGAAGACTAGCCTTACAGGCCGGGACAAG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            AsnGlnValGluGlyGluValGlnIleValSerThrAlaThrGlnThrPheLeuAlaThr
                                                                                      HCV; hepatitis C virus; single chain recombinant complex; linker; NS3 protease; NS4A cofactor; NS4A-NS3 complex; serine protease, hydrophobic domain; covalent complex; detection; inhibitor; ss.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Sequence 1998 BP; 411 A; 595 C; 569 G; 423 T; 0 other;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Conservative:
Mismatches:
Indels:
                                                          ACV NS4A-NS3 complex encoding cDNA SEQ ID NO:105
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Length:
Matches:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Disclosure; Page 166-169; 211pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                          New hepatitis C virus covalent complexes
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Gaps:
                                                                                                                                                                                                                                                                                                                                                                               Yao
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 US-09-965-594-20 (1-197) x AAX80355 (1-1998)
                                                                                                                                                                                                                                                                                                                                                                               ĸ,
                                                                                                                                                                                                                                                                                                                                                                               Weber
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1.41e-73
885.50
93.378
85.718
86.818
                                                                                                                                                                                                                                                                                                 98US-0094331
                                                                                                                                                                                                                                                                                                               97US-0067315
                             (first entry)
                                                                                                                                                                                                                                                                                                                                                                               Taremi SS,
                                                                                                                                                                                                                                                                                                                                               (SCHE ) SCHERING CORP.
                                                                                                                                                                                                                                                                                                                                                                                                           WPI; 1999-385385/32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Percent Similarity:
Best Local Similarity:
                                                                                                                                                       Hepatítis C virus.
                                                                                                                                                                                                  W09928482-A2
                                                                                                                                                                                                                                                                                                                 28-NOV-1997;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Alignment Scores:
                                                                                                                                                                                                                                                                                                 28-JUL-1998;
                                                                                                                                                                                                                                                                 24 - NOV - 1998;
                            07-SEP-1999
                                                                                                                                                                                                                                   10-JUN-1999
                                                                                                                                                                                                                                                                                                                                                                               Malcolm BA,
                                                                                                                                                                      Synthetic
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           184
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Score
a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
```

243

AAX80355 standard; cDNA; 1998

61

**4**1

other:

9

```
NS4A-NS3 complex comprising a central hydrophobic domain of native HCV NS4A-NS3 complex comprising a central hydrophobic domain of native HCV NS3 serine protease domain, where the hydrophobic domain of native HCV NS4A peptide is tethered by the linker to the amino terminus of the HCV NS3 protease domain. The present sequence encodes an example of the above complex. The covalent NS4A-NS3 complexes are useful for structural determination and determination of mode of binding of HCV inhibitors by NMR spectroscopy. They can also be used for detecting inhibitors of the protease activity, the helicase activity and the ATPASE activity of NS3. The covalent covalent covalent complexes are more soluble, stable and active than the non-
                           363
                                                       121
                                                                                 423
                                                                                                            141
                                                                                                                                      483
                                                                                                                                                                 LeuSerProArgProlleSerTyrLeuLysGlySerSerGlyGlyProLeuLeuCysPro 161
                                                                                                                                                                                            543
                                                                                                                                                                                                                      603
              ValThrArgHisAlaAspValIleProValArgArgArgGlyAspSerArgGlySerLeu
                                                                                                                            544 TCGGGGCACGCTGTGGGCCATCTTCCGGGCTGCCGTATGCACCCGGGGGGGTTGCGAAGGCG
ProLysGlyProValThrGlnMetTyrThrAsnValAspLysAspLeuValGlyTrpGln
                                                        AlaProGlnGlySerArgSerLeuThrProCysThrCysGlySerSerAspLeuTyrLeu
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     present invention describes a covalent hepatitis C virus (HCV)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                      HCV; hepatitis C virus; single chain recombinant complex; linker; NS3 protease; NS4A cofactor; NS4A-NS3 complex; serine protease; hydrophobic domain; covalent complex; detection; inhibitor; ss.
                                                                                                                                                                                                                                                                            ValAspPheIleProValGluSerLeuGluThrThrMetArgSerPro 197
                                                                                                                                                                                                                                                                                          HCV NS4A-NS3 complex encoding cDNA SEQ ID NO:109
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Disclosure; Page 179-182; 211pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              New hepatitis C virus covalent complexes
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Weber PC,
                                                                                                                                                                                                                                                                                                                                                               AAX80359 standard; cDNA; 1998
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           98WO-US24528
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     98US-0094331
97US-0067315
                                                                                                                                                                                                                                                                                                                                                                                                                    (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Taremi SS,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              (SCHE ) SCHERING CORP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  WPI; 1999-385385/32.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           24-NOV-1998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       28-JUL-1998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      WO9928482-A2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     28-NOV-1997;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Malcolm BA,
                                                                                                                                                                                                                                                                                                                                                                                                                    07-SEP-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                10-JUN-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Synthetic.
                           304
                                                        102
                                                                                                              122
                                                                                                                                                                  142
                                                                                                                                                                                            484
                                                                                                                                                                                                                                                                                                        604
                                                                                                                                                                                                                                                                                                                                                                                         AAX80359
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     The
                                                                                                                                                                                                                                                                                                                                                g
                                                                                                            οχ
                                                                                                                                  ΩQ
                                                                                                                                                                 ò
                                                                                                                                                                                      a
                                                                                                                                                                                                                      õ
                                                                                                                                                                                                                                              ద
                                                                                                                                                                                                                                                                             oy
D
                       셤
                                                     ò
```

```
141
                                                                                                                                                                                                                                                                                                                                               121
                                                                                                                                                                                                                                                                                                                  423
                                                                                                                                                                                                                                                                                                                                                                                                             162 AlaGlyHisAlaValGlyIlePheArgAlaAlaValSerThrArgGlyValAlaLysAla 181
                                                                                                                                                                                                       243
                                                                                                                                                                                                                                            303
                                                                                                                                                                                                                                                             101
                                                                                                                                                                                                                                                                               363
                                                                                                                                                                                                                                                                                                                                                                         LeuSerProArgProlleSerTyrLeuLysGlySerSerGlyGlyProLeuLeuCysPro 161
                                                                                                                       124 TCCCAACAGACGCGGGGCCTACTTGGTTGCAAGAAGACTAGCCTTACAGGCCGGGACAAG 183
                                                                                                                                                41
                                                                                                                                                                                     61
                                                                                                                                                                                                                         81
                                                                                                            21
                                                                                                                                                                                                                                                                                                ProLysGlyProValThrGlnMetTyrThrAsnValAspLysAspLeuValGlyTrpGln
                                                                                                                                                                                                                                                                       ValThrArgHisAlaAspValIleProValArgArgArgGlyAspSerArgGlySerLeu
                                                                                                                                                                                     AsnGlnValGluGlyGluValGlnIleValSerThrAlaThrGlnThrPheLeuAlaThr
                                                                                                                                                                                                                         SerIleAsnGlyValLeuTrpThrValTyrHisGlyAlaGlyThrArgThrIleAlaSer
                                                                                                                                                                                                                                                                                                                                                                                    GlySerValValIleValGlyArgIleAsnLeuSerGlyAsp-----ThrAlaTyr
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  HCV; hepatitis C virus; single chain recombinant complex; linker;
NS3 protease; NS4A cofactor; NS4A-NS3 complex; serine protease;
hydrophobic domain; covalent complex; detection; inhibitor; ss.
                                                                                                                                                                                                                                                                                                                                                                                                                                                 1998
167
16
10
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                HCV NS4A-NS3 complex encoding cDNA SEQ ID NO:104
                                             Conservative:
Mismatches:
Indels:
0
                           Length:
Matches:
569 G; 423
                                                                         Gaps:
                                                                                         US-09-965-594-20 (1-197) x AAX80359 (1-1998)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         BP.
ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         AAX80354 standard; cDNA; 1998
BP; 411 A; 595
                        2,69e-73
882.50
93,37%
85,20%
86,52%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Hepatitis C virus Synthetic.
                                             Percent Similarity:
Best Local Similarity:
Sequence 1998
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    W09928482-A2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              07-SEP-1999
                  Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      10-JUN-1999
                                                                                                            Š
                                                                                                                                                                                                                         62
                                                                                                                                                                                                                                                             82
                                                                                                                                                                                                                                                                                                                                                                         142
                                                                                                                                                                                                                                                                                                                                                                                           484
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            AAX80354;
                                                                                                                                                                                      42
                                                                Query Match:
                            ..
Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      RESULT 12
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                AAX80354
ID AAX6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                    염
                                                                                                            ŏ
                                                                                                                              g
                                                                                                                                                ò
                                                                                                                                                                  g
                                                                                                                                                                                     ò
                                                                                                                                                                                                      g
                                                                                                                                                                                                                         à
                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                              셤
                                                                                                                                                                                                                                                                                                  à
                                                                                                                                                                                                                                                                                                                 q
                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                       Q
                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                               ö
                                                                                                                                                                                                                                                                                                                                                                                                                               임
```

Page 10

us-09-965-594-20.rng

```
22 AlaginginThrArgGlyGluGinGlyCysGlnLysThrSerHisThrGlyArgAspLys 41
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Weber PC,
                                                                                                                                                                               ВP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1.58e-73
878.50
93.33%
85.64%
86.13%
                                                                                                                                                                               AAX80345 standard; cDNA; 651
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             980S-0094331.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        98WO-US24528
                                                                                                                                                                                                                                                          (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Taremi SS,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        (SCHE ) SCHERING CORP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 WPI; 1999-385385/32.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Best Local Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                Hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Percent Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                             W09928482-A2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    28-NOV-1997;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        24-NOV-1998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               28-JUL-1998;
                                                                                                                                                                                                                                                          07-SEP-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Malcolm BA,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  10-JUN-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                     Synthetic.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Query Match:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Pred. No.:
                                                                                                                                                                                               SO CCC CCC CCC X S X B X X B X X B X X B X X B X X B X X B X X B X X B X X B X X B X X B X X B X X B X X B X X B X X B X X B X X B X X B X X B X X B X X B X X B X X B X X B X X B X X B X X B X X B X X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          243
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      141
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               364 GCGCCCCCCGGGGCGCGTTCCTTGACACCATGCACCTGTGGCAGCTCAGACCTTTACTTG 423
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   162 AlaGlyHisAlaValGlyIlePheArgAlaAlaValSerThrArgGlyValAlaLysAla 181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               81
                                                                                                                                                                                                                                                                                                                NS4A NS3 complex comprising a central hydrophobic domain of native HCV NS4A NS3 complex comprising a central hydrophobic domain of native HCV NS3 serine protease domain, where the hydrophobic domain of native HCV NS4A peptide is tethered by the linker to the amino terminus of the HCV NS4 protease domain. The present sequence encodes an example of the above complex. The covalent NS4A NS3 complexes are useful for structural determination and determination of mode of binding of HCV inhibitors by NMR spectroscopy. They can also be used for detecting inhibitors of the protease activity the helicase activity and the ATPase activity of NS3. The covalent NS4A-NS3 complexes are more soluble, stable and active than the non-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   GlySerValValIleValGlyArgIleAsnLeuSerGlyAsp----ThrAlaTyr
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SerIleAsnGlyValLeuTrpThrValTyrHisGlyAlaGlyThrArgThrIleAlaSer
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      GGTTCTGTTGTTGTTGGTAGAATTATTTTATCTGGTAGTGGTAGTATCACGGCCTAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          22 AlaGlnGlnThrArgGlyGluGlnGlyCysGlnLysThrSerHisThrGlyArgAspLys
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ValThrArgHisAlaAspValIleProValArgArgArgGlyAspSerArgGlySerLeu
                                                                                                                                                                                                                                                                                                  present invention describes a covalent hepatitis C virus (HCV)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      covalent protease-peptide complexes previously available
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Sequence 1998 BP; 410 A; 596 C; 568 G; 424 T; 0 other;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1998
168
14
11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Matches:
Conservative:
Mismatches:
                                                                                                                                         ż
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Length:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Indels:
                                                                                                                                                                                                                                                            Disclosure; Page 163-166; 211pp; English.
                                                                                                                                                                                                                        covalent complexes
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Gaps:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        US-09-965-594-20 (1-197) x AAX80354 (1-1998)
                                                                                                                                         δ,
Έ
                                                                                                                                           Weber
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                3.34e-73
881.50
92.86%
85.71%
                                       98US-0094331.
97US-0067315.
    98WO-US24528
                                                                                                                                         Taremi SS,
                                                                                                                                                                                                                      hepatitis C virus
                                                                                                    (SCHE ) SCHERING CORP
                                                                                                                                                                                 WPI; 1999-385385/32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Percent Similarity:
Best Local Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Alignment Scores:
                                         28-JUL-1998;
28-NOV-1997;
  24-NOV-1998;
                                                                                                                                         Malcolm BA,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     142
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               S
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   62
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              82
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            424
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Query Match:
DB:
                                                                                                                                                                                                                        New
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      õ
      a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                엄
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
```

```
The present invention describes a covalent hepatitis C virus (HCV) NS4A-NS3 complex comprising a central hydrophobic domain of native HCV NS4A-NS3 serine protease domain, where the hydrophobic domain of native HCV NS4A peptide is rethered by the linker to the amino terminus of the HCV NS3 protease domain. The present sequence encodes an example of the above complex. The covalent NS4A-NS3 complexes are useful for structural determination and determination of mode of binding of HCV inhibitors by NNR spectroscopy. They can also be used for detecting inhibitors of the protease activity, the helicase activity and the ATPase activity of NS3. The covalent Covalent Examples are more soluble, stable and active than the non-covalent protease-peptide complexes previously available.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               5 GlySerValVallleValGlyArgIleAsnLeuSerGlyAsp------ThrAlaTyr 21
                                                                                                                                                                                                                                                                                                                                                                                                                                            HCV; hepatitis C virus; single chain recombinant complex; linker; NS3 protease; NS4A cofactor; NS4A-NS3 complex; serine protease; hydrophobic domain; covalent complex; detection; inhibitor; ss.
651
167
15
10
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              other;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                       HCV NS4A-NS3 complex encoding cDNA SEQ ID NO:95.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Sequence 651 BP; 120 A; 187 C; 200 G; 144 T; 0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Length:
Matches:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ż
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Disclosure; Page 147-148; 211pp; English
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Yao
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               New hepatitis C virus covalent complexes
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                US-09-965-594-20 (1-197) x AAX80345 (1-651)
```

```
peptide, a linker, and an HCV NS3 serine protease domain, where the
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CDNA; 1998
                                                                                                                       6.39e~73
878.50
92.86%
85.20%
86.13%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            AAX80353 standard;
                                                                                                                                                 Best Local Similarity:
                                                                                                                                         Percent Similarity
                                                                                                                Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              07-SEP-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        182
                                                                                                                                                                                                                                                                         42
                                                                                                                                                                                                                                                                                                                                                                               102
                                                                                                                                                                                                                                                                                                           62
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        604
                                                                                                                                                                                                     'n
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             AAX80353
                                                                                                                                                          Query Match:
DB:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     AAX80353
 8×88888888888×8
                                                                                                                                                                                                     ò
                                                                                                                                                                                                                    a
                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                      qq
                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                        a
                                                                                                                                                                                                                                                                                                          ð
                                                                                                                                                                                                                                                                                                                          셤
                                                                                                                                                                                                                                                                                                                                             ö
                                                                                                                                                                                                                                                                                                                                                           a
                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                              අ
                                                                                                                                                                                                                                                                                                                                                                                                                à
                                                                                                                                                                                                                                                                                                                                                                                                                               셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ****
124 TCCCAACAGACGGGGCCTACTTGGTTGCAAGAAGACTAGCCTTACAGGCCGGGACAAG 183
                                           243
                                                                              303
                                                                                                               363
                                                                                                                                121
                                                                                                                                                 423
                                                                                                                                                                  141
                                                                                                                                                                                   483
                                                                                                                                                                                                     161
                                                                                                                                                                                                                    543
                                                                                                                                                                                                                                       181
                                                                                                                                                                                                                                                       603
                                                                                              101
                         61
                                                            81
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              The present invention describes a covalent hepatitis C virus (HCV) NS4A-NS3 complex comprising a central hydrophobic domain of native HCV
                                                                                            162 AlaGlyHisAlaValGlyIlePheArgAlaAlaValSerThrArgGlyValAlaLysAla
                                                                                                                                                                                                                                               62 SerIleAsnGlyValLeuTrpThrValTyrHisGlyAlaGlyThrArgThrIleAlaSer
                                                                      244 TGCGTCAACGGCGTGTGTGGACCGTTTACCAIGGTGCTGGCTCAAAGACCTTAGCCGGC
                                                                                                                                                 364 GCGCCCCCGGGGCGCGTTCCTTGACACCATGCACCTGTGGCAGCTCAGACCTTTACTTG
                                                                                                                                                                  ValThrArgHisAlaAspValIleProValArgArgArqGlyAspSerArqGlySerLeu
                                                                                                                                                                                                    LeuSerProArgProIleSerTyrLeuLysGlySerSerGlyGlyProLeuLeuCysPro
                         AsnGlnValGluGlyGluValGlnIleValSerThrAlaThrGlnThrPheLeuAlaThr
                                                                                                                                                                                                                                                                                                                                                                                             HCV; hepatitis C virus; single chain recombinant complex; linker;
NS3 protease; NS4A cofactor; NS4A.NS3 complex; serine protease;
hydrophobic domain; covalent complex; detection; inhibitor; ss.
                                                                                                                                                                                                                                                                                 182 ValAspPheIleProValGluSerLeuGluThrThrMetArgSer 196
                                                                                                                                                                                                                                                                                                                                                                              HCV NS4A-NS3 complex encoding cDNA SEQ ID NO:108
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Disclosure; Page 176-179; 211pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            New hepatitis C virus covalent complexes
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Yao
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ۲,
                                                                                                                                                                                                                                                                                                                            BP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             describes
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Weber
                                                                                                                                                                                                                                                                                                                          AAX80358 standard; cDNA; 1998
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             98US-0094331
97US-0067315
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              98WO-US24528
                                                                                                                                                                                                                                                                                                                                                             (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SS,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SCHE ) SCHERING CORP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Taremi
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          WPI; 1999-385385/32
                                                                                                                                                                                                                                                                                                                                                                                                                                 Hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                           W09928482-A2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              24-NOV-1998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      28-NOV-1997;
                                                                                                                                                                                                                                                                                                                                                            07-SEP-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                           10-JUN-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              28-JUL-1998
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          BA,
                                                                                                                                                                                                                                                                                                                                                                                                                                         Synthetic.
                                                                                                                                                                                                    142
                                                                                                                                102
                                                                                                                                                                 122
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Malcolm
                                                                                                                                                                                                                                                                                                                  AAX80358
                                                                                                           셤
                                                                                                                                               g
                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                        a
                                                                                                                                                                                                                                                                                                                           q
                                                                                              à
                                                                                                                                ò
                                                                                                                                                                               g
                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                      ò
                         ò
                                         셤
                                                          ò
                                                                                                                                                                 à
                                                                                                                                                                                                   ò
```

```
121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         141
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               543
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              162 AlaClyHisAlaValGlyIlePheArgAlaAlaValSerThrArgGlyValAlaLysAla 181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      544 TCGGGGCACGTGTGTGTTTTTTCCGGGCTGCCGTATGCACCCGGGGGGTTGCGAAGGCG 603
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      303
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       364 GCGCCCCCCCGGGCGCGTTCTTGACACCATGCACCTGTGGCAGCTCAGACCTTTACTTG 423
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         GGTTCTGTTGTTGTTGGTAGAATTATTTTATCTGGTAGTGGTAGTATCACGGCCTAC 123
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   142 LeuSerProArgProlleSerTyrLeuLysGlySerSerGlyGlyProLeuLeuCysPro 161
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              41
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   81
hydrophobic domain of native HCV NS4A peptide is tethered by the linker to the amino terminus of the HCV NS3 protease domain. The present sequence encodes an example of the above complex. The covalent NS4A-NS3 complexes are useful for structural determination and determination of mode of binding of HCV inhibitors by NMR spectroscopy. They can also be used for detecting inhibitors of the protease activity, the helicase activity and the ATPase activity of NS3. The covalent NS4A-NS3 complexes are more soluble, stable and active than the non-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        GlyServalvalileValGlyArgileAsnLeuSerGlyAsp-----ThrAlaTyr
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  22 AlaGlnGlnThrArgGlyGluGlnGlyCysGlnLysThrSerHisThrGlyArgAspLys
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         AsnGlnValGluGlyGluValGlnIleValSerThrAlaThrGlnThrPheLeuAlaThr
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SerIleAsnGlyValLeuTrpThrValTyrHisGlyAlaGlyThrArgThrIleAlaSer
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       122 ValThrArgHisAlaAspValIleProValArgArgArgGlyAspSerArgGlySerLeu
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ValAspPheIleProValGluSerLeuGluThrThrMetArgSerPro 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      covalent protease-peptide complexes previously available.
                                                                                                                                                                                                                                         Sequence 1998 BP; 410 A; 596 C; 568 G; 424 T; 0 other;
                                                                                                                                                                                                                                                                                                               1998
167
15
11
3
                                                                                                                                                                                                                                                                                                            Length:
Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        NO: 103.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        HCV NS4A-NS3 complex encoding cDNA SEQ ID
                                                                                                                                                                                                                                                                                                                                                                                                                                   Gaps:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                US-09-965-594-20 (1-197) x AAX80358 (1-1998)
```

us-09-965-594-20.rng

```
The present invention describes a covalent hepatitis C virus (HCV) NS4A-NS3 complex comprising a central hydrophobic domain of native HCV NS4A-NS3 complex. Comprising a central hydrophobic domain of native HCV NS3 berine procease domain, where the hydrophobic domain of native HCV NS4A peptide is tethered by the linker to the amino terminus of the HCV NS3A protease domain. The present sequence encodes an example of the above complex. The covalent NS4A-NS3 complexes are useful for structural determination and determination of mode of binding of HCV inhibitors by NMR spectroscopy. They can also be used for detecting inhibitors of the protease activity, the helicase activity and the ATPRSE activity of NS3. The covalent NS4A-NS3 complexes are more soluble, stable and active than the noncovalent protease-peptide complexes previously available.
HCV; hepatitis C virus; single chain recombinant complex; linker;
                                           NS3 protease; NS4A cofactor; NS4A-NS3 complex; serine protease; hydrophobic domain; covalent complex; detection; inhibitor; ss.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Yao N;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Disclosure; Page 160-162; 211pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                New hepatitis C virus covalent complexes
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            у
Б
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Weber
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       98US-0094331.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    98WO-US24528
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Taremi SS,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (SCHE ) SCHERING CORP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        WPI; 1999-385385/32
                                                                                                                                                                                        Repatitis C virus
                                                                                                                                                                                                                                                                                                                            W09928482-A2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    24-NOV-1998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        28-NOV-1997;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Malcolm BA,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            28-JUL-1998;
                                                                                                                                                                                                                                                                                                                                                                                                                        10-JUN-1999
                                                                                                                                                                                                                                    Synthetic
NAME OF STREET O
```

```
243
                                                                                                                                                                                                                      41
                                                                                                                                                                           61
                                                                                                                                                                                                              81
                                                                                                     GlySerValValIleValGlyArglleAsnLeuSerGlyAsp....-ThrAlaTyr
                                                                                                                                                                          42 AsnGlnValGluGlyGluValGlnIleValSerThrAlaThrGlnThrPheLeuAlaThr
                                                                                                                                                                                    62 SerIleAsnGlyValLeuTrpThrValTyrHisGlyAlaGlyThrArgThrIleAlaSer
                                                                                                                                       22 AlaGlnGlnThrArgGlyGluGlnGlyCysGlnLysThrSerHisThrGlyArgAspLys
Sequence 1998 BP; 410 A; 596 C; 568 G; 424 T; 0 other;
                         1998
167
11
11
                                          Conservative:
Mismatches:
Indels:
                         Length:
Matches:
                                                                     Gaps:
                                                                                     US-09-965-594-20 (1-197) x AAX80353 (1-1998)
                         7.93e-73
877.50
92.86%
85.20%
86.03%
                                      Percent Similarity:
Best Local Similarity:
Query Match:
DB:
                 Alignment Scores:
                                                                                                      S
                                                                                                                                                                                                                                                                                    102
                                 Score:
                                                                                                     ò
                                                                                                                     g
                                                                                                                                       ò
                                                                                                                                                       g
                                                                                                                                                                      ò
                                                                                                                                                                                         g
                                                                                                                                                                                                             ò
                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                             õ
                                                                                                                                                                                                                                                                g
```

\*

```
161
                                                                                              543
                      122 ValThrArgHisAlaAspValIleProValArgArgArgGlyAspSerArgGlySerLeu 141
                                                                                                                      181
                                                                                                                               544 TCGGGCACGCTGTGGGCATCTTCCGGGCTGCCGTATGCACCCGGGGGGTTGCGAAGGCG 603
364 GCGCCCCCCGGGGCGCGTTCCTTGACACCATGCACCTGTGGCAGCTCAGACCTTTACTTG 423
                                    AlaGlyHisAlaValGlyIlePheArgAlaAlaValSerThrArgGlyValAlaLysAla
                                                                                                                                                                    182 ValAspPheIleProValGluSerLeuGluThrThrMetArgSerPro 197
                                                                                                                                                                                162
                                            g
                                                                     ò
                                                                                         a
                                                                                                                    ò
                                                                                                                                           g
                                                                                                                                                                    õ
                                                                                                                                                                                         g
                   õ
```

Search completed: August 30, 2003, 19:48:15 Job time : 188.939 secs

```
GenCore version 5.1.6
```

compugen Lta.	s_p2n model
copyright (c) 1993 - 2003 compugen Ltd.	using frame_plu
copyright (	OM protein - nucleic search, using frame_plus_p2n model
	OM protein

econds updates/sec	RSP 197			
Search time 1910.31 Seconds (without alignments) 2506.388 Million cell updates/sec	VAKAVDFIPVESLETIMRSP 197			45562784
August 30, 2003, 19:20:43 ; Search time 1910.31 Seconds (without alignments) 2506.388 Million cell update:	LSGDTA	apext 0.5 apext 0.5 apext 7.0 slext 7.0	22781392 segs, 12152238056 residues	Total number of hits satisfying chosen parameters:
August 30, 2003,	US-09-965-594-20 1020 1 MKKKGSVVIVGRINLSGDTA	BLOSUM62 Xgapop 10.0 , Xgapoxt Ygapop 10.0 , Ygapext Fgapop 6.0 , Fgapext Delop 6.0 , Delext	22781392 segs, 1	hits satisfying
Run on:	Title: Perfect score: Sequence:	Scoring table:	Searched:	Total number of

Minimum DB seq length: 0 Maximum DB seq length: 200000000

Post-processing: Minimum Match 0% Maximum Match 100% Listing first 45 summaries

Command line parameters:
-MODGL-frame+\_p2n.model -DEV-x1p
-MODGL-frame+\_p2n.model -DEV-x1p
-MODGL-frame+\_p2n.model -DEV-x1p
-DE-EGT -DEPC-SPOOL/US09965594/runat\_29082003\_151919\_28322/app\_query.fasta\_1.2872
-DB-EGT -OFHT-fastap -SUFFIX-rst -MINMARCH-0.1 -LOOPCL-0 -LOOPEXT-0
-UNITS-b1ts -START-1 -END-1 -MATRIX-blosum62 -TRANS-human40.cd1 -LIST-45
-UNITS-b1ts -START-1 -END-1 -THR\_MAX-100 -THR\_MIN-0 -ALIGN-15 -MODE-LOCAL
-OUTFRY-pto -NORM-ext -HEAPSIZE-500 -MIRIN-0 -MAXLEN-200000000
-USER-US09965594\_eCGN1\_1\_12650\_erunat\_29082003\_151919\_28322 -NCPU-6 -ICPU-3
-NO\_MANA -LAREDQUERY -NEG\_SCORES-0 -WAIT -DSPBLOCK-100 -LONGLOG
-DEV\_TIMPOUT-120 -WARN\_TIMDOUT-30 -THREADS-1 -XGRAPOP-10 -XGAPEXT-0.5 -FGAPOP-6

em\_gss\_vrt:\*
em\_gss\_fun:\*
em\_gss\_mam:\* gb\_est4:\*
gb\_est5:\*
em\_estfun:\* em\_gss\_hum:\* em\_gss\_inv:\* em\_gss\_pln:\* em\_gss\_pro:\* em\_gss\_rod:\* em\_gss\_mus: em\_esthum: em\_esthum: em\_estin: em\_estin: em\_estiv: em\_estip: em\_htc:\* gb\_esti: gb\_esti: gb\_esti: em\_estom:\* gb\_est3:\* EST:\* Database :

## 29: qb\_gss2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

D Decription	174117894	50999 CB950999 AGE	5803 BM915803	5101 BQ926101 AGE	2487 BQ892487	3727 BG089727	3988 BX238988	3244 BF863244	1699 BF304699	3021 AQ538021	3757 BE868757	50HN AL410673	5704 CC406704	2029 CC406705	2566 BM402566	182342381 182342381	1/// BUS41777	BQ894657	534 AAU36834	1249 A0620249	9264 BU299264	14.38 BU240438	755 BU365755	2542 CC380642	230B	2510 BU402910	3343 BD219343	2842 BU542842	1207 BX434207	73208 CC373208	AV835401	J9179 BX059665	55.54 BU3969.24	2000	20113 20545	57279 BM467279	74 BM553374 AG	51518 AVOR1510 W:	BTCTCOVG STCT	25317 BM825317 K-E	5938 CAS 245	361059 461059	
108	} :	4	a	m	m	0	o	0	0	æ	0	σ	on i	י עכ	2	<b>.</b> .	<b>.</b>	ີ .	• (	× 0x	m (	m) (	<b>~</b> ) (	ם תב	n =		m		m		۳.,	٦,	n c	٠.	۰, د	٦,	10	;	10	1.	10	-	
enath		1031	7	0	13	$\sim$	*	0	∞	ø	93	9	~	ю (	~ ·	o ,	٦,	<b>n</b> (	$\rightarrow$ $^{\circ}$	ν,	Н,	- 1	~ :	~ r	۰α	່ຕ	9	0	0	936	ν,	٦,	N U	, 6	3 5	7 7	· ~	; ;	15	v	<b>,</b> –	. ~	١
Ouery Match 1		₹.				•				•					٠						•	٠			•				٠	æ (				٠	٠	•			•				•
Score	:	106.5	7		2	• •	ַ ס	95.5	'n	<b>7</b> 6	94	σ,	93.5	ς,		2) (	٠,	, ,		i.	91.5	∹.	∹'	7.5	15	16	91	91	16	•			0.00	· c	· • •				;σ	0.6	06	0	
Result No.			c 0			υ v.			စ ပ	D	10		c 15	Υ ·	٦,	5	٦,	٦,	) 1	4 (	0 50 0 50	7 (	<b>N</b> (	40	25	26	~	c 28	~	30	<b>,</b>	יו רי	37	"	) ("	37	(1)	~	0 4	4	4	C 43	

## ALIGNMENTS

CB950999 1031 bp mRNA linear EST 29-APR-2003 AGENCOURT\_13445496 NIH\_MGC\_177 Mus musculus cDNA clone IMAGE:30316162 5', mRNA sequence. EST.

Mus musculus (house mouse)

Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 (bases 1 to 1031) CB950999.1 GI:30205777 CB950999 LOCUS ACCESSION VERSION KEYWORDS SOURCE ORGANISM RESULT 1 CB950999

REFERENCE

em\_gss\_phg:\*
em\_gss\_vrl:\*
gb\_gssl:\*

~

us-09-965-594-20.rst

```
160 sProAlaGlyHisAlaValGlyIlePheArgAlaAlaValSerThrArgGlyValAlaLy 180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Length:
Matches:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              US-09-965-594-20 (1-197) x BM915803 (1-1146)
                                                                                                                                                                                                                                            5', mRNA sequence.
BM915803
BM915803.1 GI:19366182
                                                                                                                                                                                                                                                                                                                             Homo sapiens (human)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     106.00
36.26%
25.82%
10.39%
                                                                                                                                                                                                                                                                                                                                                  Homo sapiens
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Best Local Similarity:
Query Match:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Percent Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Alignment Scores:
                                                                                  180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Source
                                                                                                                                                                                                                                                                                    VERSION
KEYWORDS
SOURCE
ORGANISM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       BASE COUNT
ORIGIN
                                                                                                                                                               RESULT 2
BM915803/c
                                                                                                                                                                                                                         DEFINITION
                                                                                                                                                                                                                                                                                                                                                                                                         REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
                                                                                                                                                                                                                                                                 ACCESSION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    FEATURES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Pred.
                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ð
                                                                                                                                                                                                                                                                                                                                                                         /ullerarel raddenilors/
/lab_host="DN10B (T1-phage-resistant)"
/lab_host="DN10B (T1-phage-resistant)"
/clone="Lib"NH-MCG-177"
/clone="Corgan: Ilver; Vector: pDNR-LIB; Site_1: Sfil
/clone="Corgan: Ilver; Vector: pDNR-LIB; Site_1: Sfil
/cocattatggcc]: Site_2: Sfil (ggccgctcggcc); cDNA made
by oligo off priming and directionally cloned. S' and 3'
adaptors were used in cloning as follows:
S'-AAGCAGTGATACACGCAGAGTGCCATTACGCGCGGCGGC
S'-AATCTAGAGGCGAGCGCGCGACATG-dT(30)NN-3'. Full-length
enriched library was constructed using the Clontech
Creator SMART kit and size-selected to contain the 0.5 kb
size fraction. Library created in the laboratory of M.
Brownstein (NIMH, NIH). Note: this is a NH_MGC Library.
35 a 309 c 211 g 275 t lothers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               124 ArgHisAlaAsp-ValIleProValArgArgArgGlyAspSerArgGlySerLeuLeuSe 143
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                104
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          509 ------ACACTTAACCT-CCTTGCCTGGCCTGGCATGTTGGGCAAAAGAACGTTT 558
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  143 rProArgProlleSerTyrLeuLysGlySerSerGly------GlyProLeuLeuCy 160
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Dr. Michael Brownstein
Tissue Procurement: Dr. Michael Brownstein
CDNA Library Preparation: Michael Brownstein Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: MCMIO7 row: b column: 11
High quality sequence stop: 333.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     356 CAAGIGIGCCACCIGCICACCCAC.----AAGGIICAGIAIGIGAIICAAGGCIAICCC 409
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            467 TCCGIGGCACAT-------ATCGATCGCCCTAAAGGGCCTTCACAAAA 508
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ---GlySerArgSerLeuThrProCysThrCysGlySerSerAspLeuTyrLeuValThr 123
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       559 TIGGGIICGCCCCCTIGGCCCCCCCCAAIIGGGAACCAGIGGC------ACC 606
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         49 GlnIleValSerThrAlaThrGlnThrPheLeuAlaThrSerIleAsnGlyValLeuTrp 68
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     29 GlnGlyCysGlnLysThrSerHisThrGlyArgAspLysAsnGlnValGluGlyGluVal 48
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    89 MetTyrThrAsnValAspLysAspLeuValGlyTrpGlnAlaProGln-------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ACCCATGGGGCTTGTTGTTCCCCGCCTCTCCCGTGGGCAATTACAAAACNCCCTTAACCG
NIH-MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1031
54
19
69
37
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Length:
Matches:
Conservative:
Mismatches:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Indels:
                                                                                                                                                                                                                                                                                                     1. .1031
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Gaps:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        US-09-965-594-20 (1-197) x CB950999 (1-1031)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              2.23
106.50
41.018
30.348
10.448
                                      Unpublished
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Percent Similarity:
Best Local Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     105
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Query Match:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Pred. No.:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               BASE COUNT
ORIGIN
                 TITLE
JOURNAL
COMMENT
AUTHORS
                                                                                                                                                                                                                                                                                  FEATURES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
```

```
::: ::::::: 111
1038 CTGGAGGACAGAGGT----------CTACGCGCGTAGGGGA 1003
                                                                                                                                                                      BM915803 1146 bp mRNA linear EST 12-MAR-2002 AGENCOURT_6639455 NIH_MGC_41 Homo sapiens cDNA clone IMAGE:5482056
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          22 AlaGinGlaThrArgGlyGluGlnGlyCysGlnLysThrSer-HisThrGlyArgAspLy 41
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       41 sAsnGlnValGluGlyGluValGlnIleValSerThrAlaThrGlnThrPheLeuAlaTh 61
                                                                                                                                                                                                                                                                                                                           Eukaryota: Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1146)
NH-MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC) Unpublished
                                                                                                                                                                                                                                                                                                                                                                                                                                            Contact: Robert Strausberg, Ph.D.
Email: cgapbs-rémail.nih.gov
Tissue Procurement: DCTP/DTP
CDNA Library Preparation: Rubin Laboratory
CDNA Library Preparation: Rubin Laboratory
CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCARGOT row: i column: 01
High quality sequence start: 6
High quality sequence stop: 256.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ģ
sAlavalAspPheIleProValGluSerLeuGluThrThrMetArgSerPro 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1146
47
19
59
57
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Conservative:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             윱
```

BASE COUNT 297 a 521 c 237 g 345 t 3 others ORIGIN	Alignment Scores: Pred. No.: Pred. No.: Score: Score: Score: Score: Score: 14 Bercent Similarity: 29.00% Mismatches: 66 Query Match: 10.05% Indels: 62 DB:	US-09-965-594-20 (1-197) x BQ926101 (1-1403)  QY	0y 45 GluGlyGluValGlnIleValSerThrAlaThrGlnThrPheLeuAlaThrSerIleAsn 64	104   10G1ySerArgSerLeuThrProCysThrCysG1ySerSerAspLeuTyrLeuValThrA	161 ProAlaGlyHisAlavalGlyIlePheArgalaAlavalSerThrArgGlyVal 178 161 ProAlaGlyHisAlavalGlyIlePheArgalaAlavalSerThrArgGlyVal 178 162 AGGAGAGGCGGCGTGTTTGGGCGGTGAGGACGAAGCGCGGTG 875 163 BQ892487 163 BQ892487 164 BQ892487 165 BQ892487 167 BQ892487 177
61 rSerlleAsnGlyValLeuTrpThrValTyrHisGlyAlaGlyThr	LysGlyProValThrGlmMetTy 90 :::       :::        :::        :::	107 gSerLeuThrProCysThrCysGlySerSerAspLeuTyrLeuValThrArgHisAlaAs 107 gSerLeuThrProCysThrCysGlySerSerAspLeuTyrLeuValThrArgHisAlaAs 111	0y         147 eSerTytLedLysGlySerSerGlyGlyProLeuLedCysProAladLyHisAlaVaiGl 167           0y         1   1   1   1   1   1   1   1   1   1	. o z x	CONTACT: Robert Strausberg, Ph.D.  Tissue Procurement: Mark Maconochie, Ph.D. and Nancy L. Freeman, Ph.D.  CONTACT: Robert Strausberg, Ph.D.  Email: cgapbs-r@mail.nih.gov  Tissue Procurement: Mark Maconochie, Ph.D. and Nancy L. Freeman, Ph.D.  CONA Library Preparation: ResGen, Invitrogen Corp  CONA Library Prayed by: The I.M.A.G.E. Consortium (LiML)  DNA Sequencing by: Agencourt Blossience Corporation  Clone distribution: MGC clone distribution information can be http://image.linl.gov  Plate: LiAML3798 row: j column: 07  High quality sequence stop: 101.  Account Account Consortium/Linl. at:  L. 1403  Account Account Consortium (Consortium)  Account Constructed by ResGen, Invitrogen Corp. Note: this is a

```
Contect: Robert Strausberg, Ph.D.

Email: ggapbs-r@mail.nh.gov

Enail: cgapbs-r@mail.nh.gov

Tissue Procurement: David Segal Ph.D., Herbert Morse M.D.

CDNA Library Preparation: Life Technologies, Inc.

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Washington University Genome Sequencing Center

Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov
                                                                                                                                                                                                                                                                                                                                               BG089727 629 bp mRNA linear EST 26-JAN-2001 mab90e06.x1 NCI_CGAP_SP2 Mus musculus cDNA clone IMAGE:3977578 3' similar to SW:GRAD_MOUSE P11033 GRANZYME D PRECURSOR;, mRNA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       /clone='INAGE:397758"
/tissue_type="NK cells (flow-sorted)"
/tab_nost='DH10B (Tl-resistant)"
/clone_lib="NC_CGB-5p2"
/note="Organ: spleen: Vector: pCMV-SPORT6 (Life
/note="Organ: spleen: Vector: pCMV-SPORT6 (Life
Technologies); mRNA made from flow-sorted NK cells, cDNA
made by oligo-dT priming. Directionally cloned. Average
insert size 1.5 kb. Primary library, non-amplified. cDNA
Library Preparation: David B. Krizman, Ph.D."

156 c 150 g 191 t l others
    157 oLeuLeuCysProAlaGlyHisAlaValGlyIlePheArgAlaAlaValSerThrArgGl 177
                                                                                                                                                                        817 CCTAACATCTCCTACCCCTGCGCGCGCGGGGGGGGGGCGCGGGCGCATACGGGCTCAGGG 876
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalla; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 629)
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
NTACIOGAP http://www.ncbi.nlm.nih.gov/ncicgap.
Tumor Gene Index
Unpublished
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            HisThrGlyArgAspLysAsnGlnValGluGlyGluValGlnIleValSerThrAlaThr 55
                                               --GlySerSerGlyGlyPr
                                                                                       CCACCCTCGGCCCACCGATGCGCTAAGCCTTTTACAAGCCACGGCCCGGGGCCCCCC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             629
48
23
59
51
                                                                                                                                                                                                                   177 yvalalaLysAlaValAspPhelleProValGluSer 189
                                                                                                                                                                                                                                                          877 cgrittaAAAGccccccGccritcGcccccGGccGAAGCA 913
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Conservative:
Mismatches:
Indels:
                                               -SerProArgProlleSerTyrLeuLys----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Matches:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Length:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             /organism-"Mus musculus"
/mol_type-"mRNA"
/db_xref-"taxon:10090"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               US-09-965-594-20 (1-197) x BG089727 (1-629)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Mus musculus (house mouse)
                                                                                                                                                                                                                                                                                                                                                                                                                                                       BG089727.1 GI:12572290
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             9.21
97.50
39.23%
26.52%
9.56%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Percent Similarity:
Best Local Similarity:
Query Match:
DB:
                                                                                                                                                                                                                                                                                                                                                                                                                   sedneuce
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Alignment Scores:
                                                                                                                                                                                                                                                                                                    RESULT 5
BG089727/c
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       BASE COUNT
ORIGIN
                                                                                                                                                                                                                                                                                                                                                                                                                                   ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
                                                                                                                                                                                                                                                                                                                                                                    DEFINITION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Pred. No.:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         AUTHORS
TITLE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   REFERENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    JOURNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       FEATURES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Score:
                                               ò
                                                                                   D<sub>D</sub>
                                                                                                                              ò
                                                                                                                                                                    ΩD
                                                                                                                                                                                                                   \overset{\circ}{\Delta}
                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                     142
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     430 CCAGCCACTGTCTTCATGTAATGACCCTTTCCCACCACACAGGGCACACGCATGGGAT 489
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          490 CCATTTTAATAAAGGGTGCTCTGTTAATCATGGCGCCCACGGGCCCGCTGATACTTCCA 549
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                90 TyrThrAsnValAspLysAspLeuValGlyTrpGlnAlaProGlnGlySerArgSerLeu 109
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        989
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   370 CATGGGGAAATCGCCCCCTTCAGAGACGAGGTCGTTCATGTTTTCTGAAAACATAACCG 429
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               53
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    83
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            SerGlyAspThrAlaTyrAlaGlnGlnThrArgGlyGluGlnGlyCysGlnLysThrSer 35
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   69
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         54 AlaThrGlnThrPheLeu-----AlaThrSerlleAsnGlyValLeuTrpThr
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 577 GCTGCCTGCACACCCCCCATGACCGATGTTGGGCTTATGTGGAAACGGCGAGGCGTTC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   -----ArgArgGlyAspSerArgGlySerLeuLeu------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 70 VallyrHisGlyAlaGlyThrArgThrIleAlaSerProLysGlyProValThrGlnMet
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                TITACCACATGIGACAGTGACTIT ---------TGT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      HisThrGlyArgAspLysAsnGlnValGluGlyGluValGlnIleValSerThr-----
                                                                     ဗူ
               CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can http://image.lln.gov
http://image.lln.gov
Plate: LiAM13595 row: c column: 13
High quality sequence start: 57
High quality sequence stop: 394.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              121 -LeuValThr------ArgHisAlaAspValIleProValArg----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            28 others
cDNA Library Preparation: Life Technologies, Inc.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Length:
Matches:
Conservative:
Mismatches:
Indels:
Gaps:
                                                                                                                                                                                                                                                                                                                                           /tissuc_type="sympathetic trunk"
/dev_stage="adult, 16 yr"
/lab_host="DH108"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             211 t
                                                                                                                                                                                                                                   /organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6192708"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               US-09-965-594-20 (1-197) x BQ892487 (1-1199)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             343 9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Technologies.
362 c
                                                                                                                                                                                                                                                                                                                           /sex-"male"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Percent Similarity:
Best Local Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             36
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                550
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            16
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Query Match:
DB:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             BASE COUNT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ..
9
                                                                                                                                                                                               FEATURES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ORIGIN
```

õ đ ò qq δλ g ô g ò අ à 셤 ò å Ŋ

Pred. No.:         10.7         Length:         644           Score:         97.00         Matches:         48           Percent Similarity:         41.95%         Conservative:         25           Best Local Similarity:         27.59%         Mismatches:         84           Query Match:         9.51%         Indels:         17           DB:         29         Gaps:         7	US-09-965-594-20 (1-197) x BX238988 (1-644)  QY		376 TCGAGGCGCTGTCCCCGTCACCAGTGACCAGCAGCTCATGTCGGTGGTCCAGCAG 92 nValAsplysAspLeuValGlyTrpGlnAlaProGlnGlySerArgSe ::::	265 GTCACCACCATGCCTTTAAAAAGGAGAAATTATTTGTGATTACCATGGGGAGGAGTTG 128 IlleProvalargArgGlyAspSerArgGlySerLeuLeuSerProArgPro 128 IlleProvalArgArgGryAspSerArgGlySerLeuLeuSerProArgPro 1205 TCGCACGAGGAGGCTTAGGGAGGCCTCCCCCGTGTCTCCTCCTCCTCATG 147 -IleSerTyrLeuLySGlySerSerGlyGlyProLeuLeuCysProAlaGlyHisAl	Db 145 TATTCTTTAAAGGCTTGTGGGAGGACCTTTGCCTGGATGCCCAGTCTTGGCC 89  Qy 165 aValGly1lePheArgAlaAlaValSerThrArgGlyVal 178  ::		ORGANISM Chlamydomonas reinhardtii  Bukaryota: Vyridiplantae: Chlorophyta; Chlorophyceae: Volvocales; Chlamydomonadaceae: Chlamydomonas.  REFERENCE 1 (bases 1 to 701) AUTHORS Grossman, A., Davies, J., Federspiel, N., Harris, E., Hauser, C., Lefebyre, P., McDermott, J. P., Shrager, J., Silflow, C. and Stern, D. TILLE Analyses of the Chlamydomonas reinhardtii Genome: A Model,	Unicellular System for Analyzing Gene Function and Regulation in Vascular Plants; project phase 3  JOURNAL Unpublished  COMMENT Contact: Charles Hauser  DOMB Box 91000  Duke University  Durham, NC 27708-1000  Tel: 919 613 8159  Fax: 919 613 8177  Email: chauser@duke.edu.  FEATURES  Location/Qualifiers  1. 701
Db 620 CATCCGGGTAAGGAAGGAGCACACACACATCATCCTTGTGCA 579  Qy 56 GlnThrPhbLeualaThrSerIleAsnGlyValLeuTrpThrValTyTHisGly 73  :::	Qy 74	Qy         102 AlaProGlnGlySerArgSerLeuThrProCysThrCysGlySerSerAspLeu 119           Db	SerTyr  IlePhe :::    CTTTTC	170 ArgAlaAlavalSerThrArgGlyValAlaLysAlaValAspPhelle	O N N	S GSS.  Danio rerio (zebrafish)  ISM Danio rerio (zebrafish)  ISM Danio rerio (zebrafish)  Eukaryota: Metazoa: Chordata; Craniata: Vertebrata; Euteleost  Eukaryota: Metazoa: Cyprinidae; Danio.  Cypriniformes: Cyprinidae; Danio.  CE 1 (bases 1 to 644)  RS Humphray.5.J., Huckle,E. and Durham,J.L.  Direct Submission	JOURNAL Submitted (27-JAN-2003) The Sanger Institute, Wellcome Trust Genome Campus, Hinxton, Campridgeshire, CBIO 1SA, UK. E-mail enquiries: hunquery@sanger.ac.uk Unpublished COMMENT This sequence was generated from the T7 end of BAC 283L13. 283L13 is part of the Danlokey BAC Library created by R. Plasterk and N.V. Keygene. Further details: http://www.sanger.ac.uk/Projects/D_rerio/.	FEALURES LOCALLON JUMINITIES SOURCE 1. 644 / Organism="Danio rerio" / Mol_type="genomic DNA" / Mol_type="genomic DNA" / Mol_type="genomic DNA" / Allone="DKEY-28113" / Clone="DKEY-28113" / Lissue_type="Testis" / I29 a 212 c 176 g 127 t Alignment Scores:

```
Percent Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       427
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ..
S
   ORGANISM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           BASE COUNT
                                                      REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            RESULT 9
AQ538021
                                                                                                                                                                                                                                                                                                         FEATURES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ORIGIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Pred.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       셤
                               /strain="Correction" wild type mt+ 21gr"
/db.xref="taxon:3055"
/clone=lib="Correction" cc-1690, Stress condition I,
normalized, Lambda Zap II*
normalized, Lambda Zap II*
/note="Vector: pBluescript II SK-; Site_1: EcoRI; Site_2:
XhoI; This library, constructed by John Davles and Jeffrey
McDermott, combines cDNAs from CC-1690 cells grown to
mid-log phase in TAP-N (30 min, 1hr, 4hr), TAP-S (30 min,
1hr, 4hr), TAP-P (4hr, 12hr, 24hr), No3 to NH4 (30min, 1hr,
4hr), TAP-P (4hr, 12hr, 24hr), No3 to NH4 (30min, 1hr,
4hr) and NH4 to NO3 (30min, 1hr, 4hr). PolyA mRNA was
purified from each sample, pooled and cDNA synthesized.
The cDNA was directionally cloned into lambda Zap II
(Stratagene) in the EcoRI (5') and XhoRI (3') sites.
pBluescript II SK- plasmids were excised from the lambda
ZAP clones by superinfection with ExAssist (Stratagene)
phage- The library was normalized using method 4 described
in Bonaldo et al (1996) Genome Research 6: 791-806."
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  BF304699 984 bp mRNA linear EST 21-NOV-2000 6018888252F1 NIH_MGC_17 Homo sapiens cDNA clone IMAGE:412276 5',
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              149
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              138
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 405 GAAGGGGACGGGAATICCGTTTGCGGAAAAGTGAGCGCCCCAAGGTCTGACCAAGTGCTA 464
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               150 LeuLysGlySerSerGlyGlyProLeuLeuCysProAlaGlyHisAlaValGlyIlePhe 169
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                284
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                107 ArgSerLeuThrProCysThrCysGlySerSerAspLeuTyrLeuValThrArgHisAla 126
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          404
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   524
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     170 ArgAlaAlaVal-------------SerThrArgGlyValAlaLySAla--- 181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ThrGlnMetTyrThrAsnValAspLysAspLeuValGlyTrpGlnAlaProGlnGlySer 106
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   AGAAGCCCGACCGCTGGTCTCTGGGGTCATCCGCATGCTATGCAATCTCCCGCTATCAG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              -----GlySerLeuLeuSerProArgProIleSer---Tyr
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 465 CTCCAAGGCAGCAATGGGAGCCTTTCGCGGTGTGCGGGTGCTCGTCCTCTAATGTCAGG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         GAGATCATTTGCATGTGGCTTTAGTCACCCCAAGAAGAGCCTGGGAGTGGGCATTTATAA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      525 AAAGAAACCATTTGAGTAGGAAGTCGCCGGTTACCCCCGAAGGTGAAGGTCACCTCTAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              231 ACAAAGTTACATACACGG-----AAGGACCAGCGCGCTTGGCCCCCCTTGGAGCCG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            182 ValAspPheIleProValGluSerLeuGluThrThrMetArg 195
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  'organism="Chlamydomonas reinhardtii"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       71 TyrHisGlyAlaGlyThrArgThrIleAlaSerProLys----
                                                                                                                                                                                                                                                                                                                                                                                                              701
38
22
22
63
31
                                                                                                                                                                                                                                                                                                                                                                                                            Length:
Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       US-09-965-594-20 (1-197) x BFB63244 (1-701)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          BF304699.1 GI:11251586
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             saptens (human)
                                                                                                                                                                                                                                                                                                                                                                                                            16.8
95.50
38.96%
24.68%
9.36%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                mRNA sequence.
BF304699
                                                                                                                                                                                                                                                                                                                                                                                                                                                 Percent Simllarity:
Best Local Similarity:
                                                                                                                                                                                                                                                                                                                                                                                               Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               HOMO
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            EST
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              87
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          345
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            139
                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Query Match:
DB:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               RESULT 8
BF304699/c
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   DEFINITION
                                                                                                                                                                                                                                                                                                                                         BASE COUNT
ORIGIN
                                                                                                                                                                                                                                                                                                                                                                                                                  No.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        VERSION
KEYWORDS
SOURCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ACCESSION
                                                                                                                                                                                                                                                                                                                                                                                                                                  Score:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          õ
```

```
/organism="Homo sapiens"
/wol_type="mRNA"
/wol_type="taxon:9606"
/db_xref="taxon:9606"
/clone="IMAGE-412276"
/clone_lib="Host="DH10B (phage-resistant)"
/clone_lib="MIH_MGC_17"
/note="Organ: muscle: Vector: poTB7; Site_1: EcoRI;
Site_2: XhoI; cDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGGAGGGG; Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using 2AP-CDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies)."
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         100 TrpGlnAlaProGlnGlySerArgSerLeuThr---ProCysThrCysGlySerSerAsp 118
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          159 LeuCysProAlaGlyHis-Ala------ValGlyIlePheArg-----AlaAlaVa 173
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    554
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                139 GlySerLeuLeuSerProArgProlleSerTyrLeuLysGlySerSerGlyGlyProLeu 158
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    553 GGGGGCCTCTTG------TGGGGAGACCCGA 527
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 526 TGGTGTCCAAGCTCGCGCTGCTGTACTGGAAGTCGGCACGCTCCGGTCAGGTGCAGCTTC 467
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    173 ISerThrArgGlyValAlaLySAlaValAspPheIleProValGluSerLeuGluThrTh 193
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalla; Eutheria; Primates; Catarrhlni; Hominidae; Homo.
                                                                                                                                                                                       Email: cgapbs-rémail.nih.gov
Tissue Procurement: ArCC
CDNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incytte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
Plate: LLCMI005 row: g column: 13
High quality sequence stop: 646.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1 (bases 1 to 984)
NIH-MGC http://mgc.nci.nih.gov/.
National institutes of Health, Mammalian Gene Collection (MGC)
Conpublished
Contact: Robert Strausberg, Ph.D.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             646 rescensificacescarrecesisces as Assassas actions ---
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 984
66
26
34
7
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Conservative:
Mismatches:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Length:
Matches:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Gaps:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           US-09-965-594-20 (1-197) x BF304699 (1-984)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CAGCGCCCGGGGTGCG-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 193 rMetArgSerPro 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CAAGGGGGCGCCC 415
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              26.2
95.50
42.86%
37.14%
9.36%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1. .984
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Best Local Similarity:
Query Match:
```

```
BE868757 935 bp mRNA linear EST 20-OCT-2000 601445993F1 NIH_MGC_65 Homo sapiens cDNA clone IMAGE:3850079 5',
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       /organism="Homo sapiens"
/mol_type="mxNA"
/db_xref="taxon:9606"
/clone="IMAGE:3950079"
/tissue_type="adenocarcinoma"
/tissue_t
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ----ArgSerLeuThrProCysThrCysGlySerSerAspLeuTy 120
                                                                                                                    120 rLeuValThrArgHisAlaAspValIleProValArgArgArgGlyAspSerArgGlySe 140
                                                                                                                                                                                                                                     140 rLeuLeuSerProArgProIleSerTyrLeuLysGlySer-SerGlyGlyProLeuLeuC 160
                                                                                                                                                                                                                                                                                               361
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1 MetLysLysLysGlySerValVallleValGly-----ArgIleAsnLeuSerGly 17
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Eukaryota, Metazoa; Chordata; Craniata; Vertebrata; Buteleostomi; Mammalia; Butheria; Primates; Catarrhini; Hominidae; Homo. 1 (bases 1 to 935) NIH-MGC http://mgc.nci.nih.gov/. National Institutes of Health, Mammalian Gene Collection (MGC)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        cDNA Library Preparation: Life Technologies, Inc.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM9568 row: i column: 24
High quality sequence stop: 557.
Location/oualifiers
1. 935
                                                                                                                                                        311 TCTCCCTGGCCCAGCCCTGGACCATATCTGAAAGCTGTGAAGTGGAGGC-----T
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            935
117
128
6
128
6
                                                                                                                                                                                                                                                                                                                                                         160 ysproalaglyHisalaValGlyIlePheArgalaAla 172
                                                                                                                                                                                                                                                                                                                                                                                                             362 GCGGCGGGGAGCTGATGCAGCTGTTCTCACAGGGGCT 399
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Conservative:
Mismatches:
                                                             230 ACTIGCIGCAIGGGGAGGAGGICTGIACAIAGAIGI-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        145 t
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Length:
Matches:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Contact: Robert Strausberg, Ph.D.
Email: cgapbs-rémail.nih.gov
Tissue Procurement: ATCC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      US-09-965-594-20 (1-197) x BE868757 (1-935)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        261 9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     BE868757.1 GI:10317533
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Technologies.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Homo sapiens (human)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            34.5
94.00
42.66%
30.77%
9.22%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        245 C
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              mRNA sequence.
BE868757
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Homo sapiens
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Unpublished
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Percent Similarity:
Best Local Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ĸ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        284
      106 r----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Query Match:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        VERSION
KEYWORDS
SOURCE
ORGANISM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     source
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Pred. No.:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               DEFINITION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           COUNT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             REFERENCE
AUTHORS
TITLE
JOURNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                             RESULT 10
BE868757
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ACCESSION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      FEATURES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              COMMENT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        BASE CO
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                             g
                                                                                                                                                              g
                                                                                                                                                                                                                                     à
                                                                                                                                                                                                                                                                                            đ
                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                       5
                                                                                                                       ò
AQ538021 560 bp DNA linear GSS 18-MAY-1999 RPCI-11-32014.IJ RPCI-11 Homo sapiens genomic clone RPCI-11-32014,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              172
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      87 ThrGlnMetTyrThrAsn-ValAspLysAspLeuValGlyTrpGlnAlaProGlnGlySe 106
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Email: the@etigr.org
Clones are derived from the human BAC library RPCI-11. For BAC
Clones are derived from the human BAC library and a valuability, please contact Pieter de Jong
(pieter@dejong.med.buffalo.edu). Clones may be purchased from
BACPAC Resources (http://bacpac.med.buffalo.edu/ordering) or from
BACPAC Resources (info@resgen.com). BAC end search page:
http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           86 TICIGCCAGICICTGICALCCACIGGCCAAGAG-------------------------ATIGCAAA 127
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   46
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            86
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          26 CAAIGCTGTGAGAACCCAGTGGATGGACACCCTGGAAGGGCTATGAGATCATAGGAGGGC 85
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 47 GluvalGlnIleValSerThrAlaThrGlnThrPheLeuAlaThrSerIleAsnGlyVal 66
                                                                                                                                                                                                                        Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalla; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 560)
                                                                                                                                                                                                                                                                                                                          Zhao, S., Adams, M.D., Nierman, W., Malek, J., de Jong, P. and Venter
                                                                                                                                                                                                                                                                                                                                                                                    Use of BAC End Sequences from Library RPCI-11 for Sequence-Ready
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   /clone_lib-"RPCT-11*
/note_vector: pBACe3.6; Site_1: ECORI; Site_2: ECORI;
RPCI11 Human Male BAC Library*
125 c 181 g 124 t
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         128 ATGTGG------CACGGATTTGGAGCCAGGATTCCAGCTGGCTCTAAAGGG-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                GlnGlyCysGlnLys-----ThrSerHisThrGlyArgAspLysAsnGlnValGluGly
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        67 LeuTrpThrValTyrHisGlyAlaGlyThrArgThrIleAlaSerProLysGlyProVal
                                                                                                                                                                                                                                                                                                                                                                                                                                     Unpublished
Other GSSs: RPCI-11-32014.TV
Other GSSs: RPCI-11-32014.TV
Contact: Shaying Zhao, William Nierman, Mark Adams
Contact: Shaying Zhao, William Nierman, Mark Adams
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850
Far: 301 838 0208
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               560
112
138
38
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Length:
Matches:
Conservative:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Mismatches:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          /cell_type="Lymphocytes"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  /organism-"Homo sapiens"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 /mol_type="genomic DNA'/db_xref="GDB:7622691"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        /db_xref="taxon:9606"
/clone="RPCI-11-32014"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         US-09-965-594-20 (1-197) x AQ538021 (1-560)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Location/Qualiflers
1. .560
                                                          genomic survey sequence
                                                                                                                 A0538021.1 GI:4849711
                                                                                                                                                                           Homo sapiens (human)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               17.7
94.00
37.66%
29.87%
9.22%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Class: BAC ends
                                                                                                                                                                                                                                                                                                                                                                                                                   Map Building
                                                                                                                                                                                                         Homo sapiens
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Best Local Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     đ
                                                                                           AQ53802]
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Percent Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Alignment Scores:
                                                                                                                                               GSS.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   29
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Query Match:
DB:
                                                                                                                    VERSION
KEYWORDS
SOURCE
ORGANISM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                source
   LOCUS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ..
Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 BASE COUNT
                                                                                                                                                                                                                                                                                                                                                                                                                                        JOURNAL COMMENT
                                                                                        ACCESSION
                                                                                                                                                                                                                                                                                               REFERENCE
                                                                                                                                                                                                                                                                                                                             AUTHORS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           FEATURES
                                                                                                                                                                                                                                                                                                                                                                                    TITLE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ORIGIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Score:
```

Š g ò g õ g ò 셤 us-09-965-594-20.rst

S kb were prepared and both e keywords for description of the other extremity of this is source   Location/Qualifiers   1.1062   Location/Qualifiers   1.1062   Location/Qualifiers   Location/Quali	BASE COUNT 274 a 264 c 242 g ORIGIN Alignment Scores:	Pred. No.: 40.8 Le Score: 94.00 Ma Percent Similarity: 36.10% CC Best Local Similarity: 24.39% Mi Query Match: 9.22% In DB: 29.22% Ga	QY             32 GILLySThrSerHisThr                Db             447 CAAAAGACTACCCACTCTATCGATAA               QY             49 GINII eValSerThrAlaThrGlnThr               QY             69 ThrValTyr               QY             69 ThrValTyr               QY             69 ThrValTyr               QY             69 GCACTCTAAGACGAGACACCCCAGGG               QY             11
18 AspthralaTyralaGlnGlnThrArgGlyGluGlnGlyCysGlnLysThrSerHisThr 37		118 AspleuTyrLeuValThrArgHisAlaAspVallIoProvalArgArgArgGlyAspSer 137 :::	CNSOGONN  TIEGE Of Clone AWORAGOGEGIS of library AWORA from strain CLIB 89 of MITCHE ALALOGY.  MITCHER INDOLYTICA. genomic survey sequence.  ALALOGOS.1 GI:12179275  ALALOGOS.2  ALALOGOS.3  TATOWIA IIPOLYLICA.  ALALOGOS.3  TATOWIA IIPOLYLICA.  ASCOMYCOTE'S SACCHALOWYCOLINA: SACCHAROWYCETES;  SACCHAROWYCELMES: DIPOLAGOSACCEACH, TATIGUERAVE, T. Blandin, G.,  SUCIECT, J. L., Alaje, M., Artiguenave, F., Blandin, G.,  SOUCHE, J.L., Alaje, M., Artiguenave, F., Elpindia, S.,  Genomic Fubuhara, M., Bon, E., Brottler, P., Casaregola, S.,  Genomic Exploration of the hemisaccomycetous yeasts: 1. A set of yeast species for molecular evolution studies  SEBS Lett. 487 (1), 3-12 (2000)  2058471  11152876  2 (bases 1 to 1062)  Casaregola, S., Neuvegilse, C., Lepingle, A., Bon, E., Feynerol, C.,  Artiguenave, F., Wincker, P. and Galllardin, C.  Genomic exploration of the hemisaccomycetous yeasts: 17. Yarrowia ilpolytica  REBS Lett. 487 (1), 3-12 (2000)  2058471  11152892  Casaregola, S., Neuvegilse, C., Lepingle, A., Bon, E., Feynerol, C.,  Genomic exploration of the hemisaccomycetous yeasts: 17. Yarrowia ilpolytica  FEBS Lett. 487 (1), 95-100 (2000)  Direct Submission  Submission  Direct Submission  Direct Submission  Direct Submission  Sequefiedenoscope. cns. fr - Web: www genoscope. cns. fr)  This GSS is part of a tandom genomic sequencing program of thirteer: yeast species: Saccharomyces bayanus var. uvarum, Saccharomyces elegenoscope. cns. fr - Web: www genoscope.  Z rue Gaston Gremieux, CP 5706, 91037 EVRY cedex, FRANCE. (E-mail: France Submission  Direct Submission  Sequefiedenoscope. cns. fr - Web: www genoscope. cns. fr)  This GSS is part of a tandom genomic sequencing program of thirteer: yeast species: Saccharomyces bayanus var. uvarum, Saccharomyces elegenoscope. cns. fr - Web: www genoscope.  Z rue Gaston Gremieux, CP 5706, 91037 EVRY cedex, FRANCE. (E-mail: Sected Genoscope. cns. fr - Web: www genoscope.  Z rue Gaston Gremieux, CP 5706, 91037 EVRX cedex and conting angusta. Patch of a tandom genomic sequencing prog
	oy g	oy oy ob	RESULT 11 COCOS DEFINITION ACCESSION VERSION VERSION VERSION VERSION VERSION TETWORS SOUGANISM AUTHORS TITLE JOURNAL MEDLINE PUBMED REFERENCE AUTHORS TITLE JOURNAL MEDLINE PUBMED REFERENCE AUTHORS TITLE JOURNAL MEDLINE FUBMED REFERENCE AUTHORS TITLE JOURNAL COMMENT

```
GAATCATAGAGCACCAGAAAGCTGCTCTAGGAATC 1013
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       AAGTCGCCACGGCTCTGACTCTGGTTT 848
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CCGACGGGCTCAGAATTGCCACAGACGAAGATCGA 608
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       alThrArgHisAlaAspValIleProValArgArg 133
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 --GlyGlyProLeuLeuCysProAlaGlyHis--- 164
                                                                                                                                                                                                                                                                                                                ----SerLeuThrProCys-----Thr 113
                                                                                                                                                                                                                                                                                                                                           hrPheLeuAlaThrSerIleAsnGlyValLeuTrp 68
                                                                                                                                    Saccharomyces cerevisiae ORF YKR093w [
nsporter ]*
rimental
                                                                                                                                                                                                                                                                                                     lyArgAspLysAsnGlnValGlu---GlyGluVal 48
extremities were sequenced. See this sequence and for the sequence of insert.
                                                                                                                                                                   2 others
                                                                                                                                                                                                           Length:
Matches:
Conservative:
Mismatches:
Indels:
                                                      lipolytica"
DNA"
                                                                                                                                                                    280 t
                                                                                     ...
```

245

```
CC406705 789 bp DNA linear GSS 19-MAY-2003 PUHKD12TD ZM_0.6_1.0_KB Zea mays genomic clone ZMABTa469B24, genomic survey sequence. CC406705 CC406705 GI:30886795 GSS.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Tracheophyta; Spermatophyta; Tracheophyta; Spermatophyta; Paulcoldeae; Paulcoldeae; Andropogoneae; Zea.

1 (bases 1 to 789)
Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T., Resnick Maize Genomics Consortium
Unpublished
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     .....1lePro 186
                                                                                              158 LeuLeuCysProAlaGlyHisAlaValGlyIlePheArgAlaAlaValSerThrArgGly 177
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  /organism="Zea mays"
/mol_type="genomic DNA"
/strain="B3"
/strain="B3"
/db_xref="taxon:4577"
/clone="ZMMETa469B24"
/clone="Lib="ZM_0.6_1.0_KB"
/note="Vector: pCR4-TOPO; Site_1: EcoRI; 0.6-1.0 kb high
Cor selected genomic DNA library"
a 220 c 203 g 156 t
                         ArgGlySerLeuLeuSerProArgProIleSerTyrLeuLysGlySerSerGlyGlyPro 157
                                                                                                                                                                                                        244 ------GCCATTGACATCAACAAGGAGAAGTTCACCATTGAGGCATATGCAATTCCC 194
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       48 ValGlnIleValSerThrAlaThrGlnThrPheLeuAlaThrSerIleAsnGlyVal--- 66
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                29 GlnGlyCysGlnLysThrSerHisThrGlyArgAspLysAsnGlnVal --- GluGlyGlu 47
                                                                                                                     9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ::::::
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              789
49
24
67
67
10
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Matches:
Conservative:
Mismatches:
                                                                                                                                                                    178 ValAlaLysAlaValAspPhe-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ::
::
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Length:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               US-09-965-594-20 (1-197) x CC406705 (1-789)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Class: sheared ends.
Location/Qualifiers
1. .789
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Email: whitelaw@tigr.org
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Other_GSSs: PUHKD12TB
Contact: Cathy Whitelaw
                                                                                                                                                                                                                                                                        193 ITGGATACATTIGAG 179
                                                                                                                                                                                                                                             187 ValGluSerLeuGlu 191
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 93.50
39.46%
26.49%
9.17%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Seq primer: TF
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Percent Similarity:
Best Local Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Zea mays
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         TIGR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Query Match:
                                                                                                                                                                                                                                                                                                                                                                                                                             VERSION
KEYWORDS
SOURCE
ORGANISM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Pred. No.:
                                                                                                                                                                                                                                                                                                                                                                          DEFINITION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              TITLE
JOURNAL
COMMENT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      BASE COUNT
ORIGIN
                                                                                                                                                                                                                                                                                                                  RESULT 13
CC406705
LOCUS
                                                                                                                                                                                                                                                                                                                                                                                                             ACCESSION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            REFERENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             AUTHORS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     FEATURES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Score:
                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                  g
                                                                                                                             g
                                                                                                                                                                  ò
                                                                                                                                                                                                        g
                                                                                                                                                                                                                                             à
                                                                                              οy
                   PUBKD12TB ZM_0.6_1.0_KB Zea mays genomic clone ZMABTa469B24, COAAAAAAA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           550
                                                                                                                                                                                 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panitooldeae; Andropogoneae; Zea.

( bases 1 to 772)
Whitelaw, C.A., Quackenbush, J., Van Aken, S., Utterback, T., Resnick, A., Fraser, C.M., Yuan, Y., San Miguel, P., Ma, J. and Bennetzen, J. Unpublished
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     :::|||
549 ATCTGGAGG------TCACAGACTTTGCTGACGATGAGATCAAGGGTG 508
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     87 ThrGlnMetTyrThrAsnValAspLysAspLeuValGlyTrp------GlnAlaPro 103
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ACACAGATGATGACCAGG-----AAGAGCAGGACCCAGTGGTTTCCCTACATGCAATAA 454
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             GlnGlySerArgSerLeuThrProCysThrCysGlySerSerAspLeu------ 119
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               453 CTGGGATTAGAAGGGAGGACACCATGCAGCTGCGGGTAGTCCTCAACGGTCAGGAGCTGC 394
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  -----TyrLeuValThrArgHisAlaAspValIleProValArgArgArgGlyAspSer 137
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         393 TGGCCCTACTTGACACGGGTTCAACATAACTTCATCAACTGCAAGGCGGCA-CAACAG 335
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         /db_xref-"taxon:4577"
/clone="ZMGTa469B24"
/clone=lib="ZM_0.6_1.0_KB"
/clone_lib="ZM_0.6_1.0_KB"
/note="Vector: pCR4-TOPO; Site_1: EcoRI; 0.6-1.0 kb high
cor selected genomic DNA library"
1 198 c 216 g 208 t
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                663 CAGTCTCGTTCAAGGGCCTCACCCCGCCGAGATGGCGGAACGACGTAAACAAGGGCTAT 604
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ValGlnIleValSerThrAlaThrGlnThrPheLeuAlaThrSerIleAsnGlyVal--- 66
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              86
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           29 GlnGlyCysGlnLysThrSerHisThrGlyArgAspLysAsnGlnVal---GluGlyGlu 47
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     GTTACAATTGTGACG-----AGCCCTATGTACGAGGCCACCGCTGTCCCAGGCTATTCT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            LeuTrpThrValTyrHisGlyAlaGlyThrArgThrIleAlaSerProLysGlyProVal
                                                                                                                                                                                                                                                                                                                                                                                      9712 Medical Center Drive, Rockville, MD 20850, USA 121: 301-888-5843 Fax: 301-888-5043 Email: whitelaw@tigr.org Seq primer: TR Class: sheared ends. Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             772
49
24
67
67
10
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Length:
Matches:
Conservative:
Mismatches:

    772 Lorganism-"Zea mays" /mol_type-"genomic DNA" /strain-"B73"

                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Gaps:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      US-09-965-594-20 (1-197) x CC406704 (1-772)
                                                                                                                                                                                                                                                                                                                                    Other_GSSs: PUHKD12TD
Contact: Cathy Whitelaw
                                                                                                              CC406704.1 GI:30886794
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         30.1
93.50
39.46%
26.49%
9.17%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Best Local Similarity:
                                                                                                                                                  Zea mays
                                                                                            CC406704
                                                                                                                                                                       tea mays
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       150
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Percent Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                          TIGR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   48
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       603
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            67
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        207
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             104
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Query Match:
DB:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         source
                   CC406704/c
LOCUS
                                                       DEFINITION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Pred. No.:
                                                                                                                                                                    ORGANISM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    BASE COUNT
                                                                                                                                                                                                                                                                                             TITLE
JOURNAL
COMMENT
                                                                                            ACCESSION
                                                                                                                                                                                                                                                            AUTHORS
RESULT 12
                                                                                                              VERSION
KEYWORDS
SOURCE
                                                                                                                                                                                                                                             REFERENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    FEATURES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ORIGIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Score:
```

셤

ò

à a õ g ò a õ 8

ò

```
/organism="Sorghum bicolor"
/wol_type="genomic DNA"
/wol_type="genomic DNA"
/dol_type="genomic DNA"
/dol_type="genomic DNA"
/dlone="ic83bil"
/lab_host="JM107 or DH5a"
/clone=ic83bil"
/clone=ilb="wdG-SbicolorF (JM107 adapted methyl filtered)"
/clone=ilb="wdG-SbicolorF (JM107 adapted methyl filtered)"
/clone=ilb="wdG-SbicolorF (JM107 adapted methyl filtered)"
/note="Site="iSte="iSte="iSte="iSte="iSte="iSte="iSte="iSte="iSte="iSte="iSte="iSte="iSte="iSte="iSte="iSte="iSte="iSte="iSte="iSte="iSte="iSte="iSte"iste"
/note="Site="iSte="iSte="iSte="iSte="iSte="iSte="iSte="iSte="iSte="iSte="iSte="iSte="iSte="iSte="iSte="iSte="iSte="iSte="iSte="iSte="iSte="iSte="iSte="iSte="iSte="iSte="iSte="iSte="iSte="iSte="iSte="iSte="iSte="iSte="iSte="iSte="iSte="iSte="iSte="iSte="iSte="iSte="iSte="iSte="iSte="iSte="iSte="iSte="iSte="iSte="iSte="iSte="iSte="iSte="iSte="iSte="iSte="iSte="iSte="iSte="iSte="iSte="iSte="iSte="iSte="iSte="iSte="iSte="iSte="iSte="iSte="iSte="iSte="iSte="iSte="iSte="iSte="iSte="iSte="iSte="iSte="iSte="iSte="iSte="iSte="iSte="iSte="iSte="iSte="iSte="iSte="iSte="iSte="iSte="iSte="iSte="iSte="iSte="iSte="iSte="iSte="iSte="iSte="iSte="iSte="iSte="iSte="iSte="iSte="iSte="iSte="iSte="iSte="iSte="iSte="iSte="iSte="iSte="iSte="iSte="iSte="iSte="iSte="iSte="iSte="iSte="iSte="iSte="iSte="iSte="iSte="iSte="iSte="iSte="iSte="iSte="iSte="iSte="iSte="iSte="iSte="iSte="iSte="iSte="iSte="iSte="iSte="iSte="iSte="iSte="iSte="iSte="iSte="iSte="iSte="iSte="iSte="iSte="iSte="iSte="iSte="iSte="iSte="iSte="iSte="iSte="iSte="iSte="iSte="iSte="iSte="iSte="iSte="iSte="iSte="iSte="iSte="iSte="iSte="iSte="iSte="iSte="iSte="iSte="iSte="iSte="iSte="iSte="iSte="iSte="iSte="iSte="iSte="iSte="iSte="iSte="iSte="iSte="iSte="iSte="iSte="iSte="iSte="iSte="iSte="iSte="iSte="iSte="iSte="iSte="iSte="iSte="iSte="iSte="iSte="iSte="iSte="iSte="iSte="iSte="iSte="iSte="iSte="iSte="iSte="iSte="iSte="iSte="iSte="iSte="iSte="iSte="iSte="iSte="iSte="iSte="iSte="iSte="iSte="iSte="iSte="iSte="iSte="iSte="iSte="iSte="iSte="iSte="iSte="iSte="iSte="iSte="iSte=
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        119 LeuTyrLeuValThrArgHisAlaAspValIleProValArgArgArgGlyAsp----- 136
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             150 LeuLysGly------LeuCy 160
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    -----CysGlySerSerAsp 118
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                261
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                260 GTGGAGAGGGCAGGACGACGACGA-------CGACGACGGGGTGGAGACCCG 216
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              -----SerArgGlySerLeuLeuSerProArgProIleSerTyr 149
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    160 sProAlaGlyHisAlaValGlyIlePheArgAlaAlaValSerThrArgGlyVal---- 178
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                GCCCATGCCGTGCGGGGTGTTTTTTGAACGTACCGCACGGCGCGGGGAGGACGACAGGC 441
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      GIGGCCGTGGCCCGCCGCCCCCCCCGCGTGGGAATTATGGCCGCAATAAACTCGGCAC 381
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         96
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   96
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   AlaThrSerIleAsnGlyValLeuTrpThrValTyrHisGlyAlaGlyThrArgThr ---
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ----IleAlaSerProLysGlyProValThrGlnMetTyrThrAsnValAspLysAsp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 LeuValGlyTrpGlnAlaProGlnGlySerArg---SerLeuThrProCysThr----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            320 CACCCAGTGCCCCCGGCGGCCGCCGTACGGGGACAGGGAAACTGGTGGTGAGGATCAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   701
48
14
56
51
8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Length:
Matches:
Conservative:
Mismatches:
Indels:
Tel: 516 367 8884
Fax: 516 367 8874
Email: mccombleecsh.org
mate: ic83 row: b column: 11
Seq primer: -21M13UnivFwd
Class: shotgun
High quality sequence stop: 701.
Location/Qualifiers
i: 701
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          -AlaLysAlaValAspPheIlePro 186
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                US-09-965-594-20 (1-197) x BZ342381 (1-701)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  47 rececrecerrecreserres 23
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   29.8
93.00
36.69%
28.40%
9.12%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Percent Similarity:
Best Local Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      108
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      440
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   500
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           79
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              97
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 114
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           137
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Query Match:
                                                                                                                                                                                                                                                        source
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Preď. No.:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      BASE COUNT
                                                                                                                                                                                                                      FEATURES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ORIGIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
```

Search completed: August 31, 2003, 04:27:46 Job time : 1915.31 secs

```
Modified hepatitis C virus (HCV) NS3 protease comprising at least 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Wittekind M, Weinheimer S,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         WPI; 2000-465976/40.
N-PSDB; AAA73334.

        $\times$ $\times
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                WO200040707-A1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       13-JUL-2000.
872.5
868.5
868.5
868.5
868.5
864.5
864.5
864.5
861.5
861.5
                                                                                                                                                                                                                                                                                                                                                                                                  857.5
857.5
857.5
856.5
853.5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           AAB15225;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               RESULT 1
  Hepatitis C virus
                                                                                                                              August 30, 2003, 17:42:58 ; Search time 44.6227 Seconds (without alignments) 700.745 Million cell updates/sec
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1: /SIDSI/gcgdata/geneseq/geneseqp-embl/AA1980.DAT:*
2: /SIDSI/gcgdata/geneseq/geneseqp-embl/AA1981.DAT:*
3: /SIDSI/gcgdata/geneseq-embl/AA1981.DAT:*
3: /SIDSI/gcgdata/geneseq-embl/AA1983.DAT:*
4: /SIDSI/gcgdata/geneseq-embl/AA1983.DAT:*
5: /SIDSI/gcgdata/geneseq-embl/AA1983.DAT:*
6: /SIDSI/gcgdata/geneseq-embl/AA1985.DAT:*
7: /SIDSI/gcgdata/geneseq-embl/AA1986.DAT:*
8: /SIDSI/gcgdata/geneseq-embl/AA1987.DAT:*
8: /SIDSI/gcgdata/geneseq-embl/AA1980.DAT:*
8: /SIDSI/gcgdata/geneseq-embl/AA1991.DAT:*
113: /SIDSI/gcgdata/geneseq-embl/AA1991.DAT:*
114: /SIDSI/gcgdata/geneseq-embl/AA1992.DAT:*
115: /SIDSI/gcgdata/geneseq-embl/AA1993.DAT:*
116: /SIDSI/gcgdata/geneseq-embl/AA1994.DAT:*
117: /SIDSI/gcgdata/geneseq-geneseq-embl/AA1994.DAT:*
118: /SIDSI/gcgdata/geneseq-geneseq-embl/AA1994.DAT:*
119: /SIDSI/gcgdata/geneseq-geneseq-embl/AA1995.DAT:*
120: /SIDSI/gcgdata/geneseq-geneseq-embl/AA1995.DAT:*
121: /SIDSI/gcgdata/geneseq-geneseq-embl/AA1995.DAT:*
122: /SIDSI/gcgdata/geneseq-geneseq-embl/AA1995.DAT:*
123: /SIDSI/gcgdata/geneseq-geneseq-embl/AA1995.DAT:*
124: /SIDSI/gcgdata/geneseq-geneseq-embl/AA1995.DAT:*
125: /SIDSI/gcgdata/geneseq-geneseq-embl/AA1995.DAT:*
126: /SIDSI/gcgdata/geneseq-geneseq-embl/AA1995.DAT:*
127: /SIDSI/gcgdata/geneseq-geneseq-embl/AA1995.DAT:*
128: /SIDSI/gcgdata/geneseq-geneseq-embl/AA1995.DAT:*
129: /SIDSI/gcgdata/geneseq-geneseq-embl/AA1995.DAT:*
120: /SIDSI/gcgdata/geneseq-geneseq-embl/AA1995.DAT:*
120: /SIDSI/gcgdata/geneseq-geneseq-embl/AA1995.DAT:*
120: /SIDSI/gcgdata/geneseq-geneseq-embl/AA1995.DAT:*
120: /SIDSI/gcgdata/geneseq-geneseq-embl/AA1999.DAT:*
                                                                                                                                                                                                                                    1016
1 MKKRGSVVIVGRINLSGDTA.....VAKAVDFIPVESLETTMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Description
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA2002.DAT:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA2003.DAT:*
                                                                                                                                                                                                                                                                                                                                                                                              1107863
                GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.
                                                                                                                                                                                                                                                                                                                                                                                              Total number of hits satisfying chosen parameters:
                                                                                                                                                                                                                                                                                                                                                        1107863 segs, 158726573 residues
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SUMMARIES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   AAB15224
AAB15224
AAB15223
AAB15222
AAB15221
AAB15220
AAB15220
AAB15220
AAB15220
                                                                                                 OM protein - protein search, using sw model
                                                                                                                                                                                                                                                                                                BLOSUM62
Gapop 10.0 , Gapext 0.5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     A_Geneseq_19Jun03:*
1: /srr:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ü
                                                                                                                                                                                                                                                                                                                                                                                                                                   Minimum DB seq length: 0
Maximum DB seq length: 200000000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         US-09-965-594-22
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  DB
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        $
Query
Match Length D
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       197
197
197
197
195
195
665
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        100.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Score
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1010
995
980
980
963
929
929
875.5
                                                                                                                                                                                                                                                                                                Scoring table:
                                                                                                                                                                                                                                      Perfect score:
                                                                                                                                                                                                                                                                                                                                                          searched:
                                                                                                                                                                                                                                                             Sequence:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Database
                                                                                                                                       o
:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          01 m 4 m 6 m 6 m 6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Result
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Š
                                                                                                                                       Run
```

```
HCV NS4A-NS3 compl
HCV NS4A-NS3 
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Sequence encoded in Protein sequence or Peptide encoded by Sequence encoded in Sequence encoded in Peptide encoded by
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Hepatitis; NS3 protease; viral replication; chronic liver disease;
liver failure; liver cancer; mutant; mutein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Goldfarb
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Hepatitis C virus NS4A-NS3 fusion protease #7
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ALIGNMENTS
                                                                                                                                                                                                                                                                                                                                                                                AAY17881
AAY17885
AAY17888
                                                                                                                                                                                                                                                                                                                                                                                                                                           AAW93482
AAY44728
AAW77397
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     AAP92050
AAP90288
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Zhang Y,
                                                                                                                                                                                                                                                                                                                         886
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ABP71460
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    AAB15225 standard; protein; 197 AA
                                                                                                                                                                                                                           AAY17
                                                                                                                                                                                                                                                                                   AAY1
                                                                                                                                                                                                                                                                                                                                                AY.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               (BRIM ) BRISTOL-MYERS SQUIBB
    06-JAN-2000; 2000WO-US00345.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     08-JAN-1999; 99US-0115271
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  19-DEC-2000 (first entry)
    3011
3011
3011
3011
687
1766
1766
2261
2201
2436
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Hepatitis C virus
Synthetic.
```

~

```
The present sequence is a mutated version of a fusion protein created using the Hepatitis C virus (HCV) NG3 and NS4A protease enzymes. These proteins are both essential for the replication of the virus, acting to cleave its replicative proteins from the polyprotein produced from the HCV genome. Inhibitors of the two proteins should be effective as antiviral treatments of HCV infection. This is useful as HCV can lead to chronic liver disease such as cirrhosis, liver failure and liver cancer. The present invention concerns a number of NS3 mutants and NS3-NS4A fusion proteins which can be used to identify inhibitors of this type, as well as enabling structural studies of the protease and protease:inhibitor complexes. This sequence contains the alpha-helix0-7
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  LVTRHADVIPVRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVSTRGVAK 180
substitution of a hydrophobic alpha-helix 0 amino acid to a hydrophilic amino acid, useful for screening inhibitors that may treat hepatitis C
                                                                                                                                                                                                                                                                                                                                                                                                        1 MKKKGSVVIVGRINLSGDTAYAQQTRGEQGTQRTSHTGRDKNQVEGEVQIVSTATQTFLA
                                                                                                                                                                                                                                                                                                                                                                                      1 MKKKGSVVIVGRINLSGDTAYAQQTRGEQGTQKTSHTGRDKNQVEGEVQIVSTATQTFLA
                                                                                                                                                                                                                                                                                                                                                                                                                                                  61 TSINGVLWTVYHGAGTRTIASPKGPVTQMYINVDKDLVGWQAPQGSRSLTPCTCGSSDLY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Hepatitis: NS3 protease; viral replication; chronic liver disease;
                                                                                                                                                                                                                                                                                                                          100.0%; Score 1016; DB 21; Length 197; 100.0%; Pred. No. 2.1e-98;
                                                                                                                                                                                                                                                                                                                                                         Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Goldfarb V;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Hepatitis C virus NS4A-NS3 fusion protease #6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         liver failure; liver cancer; mutant; mutein
                                                                                                                                                                                                                                                                                                                                                         Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Zhang Y,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Į
                                                               Claim 23; Fig 17; 66pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            (BRIM ) BRISTOL-MYERS SQUIBB CO.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 AAB15224 Standard; protein; 197
                                                                                                                                                                                                                                                                                                                                                         ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          AVDFIPVESLETTMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        AVDFIPVESLETIMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Weinheimer S,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             99US-0115271
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                06-JAN-2000; 2000WO-US00345
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (first entry)
                                                                                                                                                                                                                                                                                                                                                         Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        WPI; 2000-465976/40.
                                                                                                                                                                                                                                                                                                                                       al Similarity
197; Conserv
                                                                                                                                                                                                                                                                                           197 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        N-PSDB; AAA73333
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  WO200040707-A1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Wittekind M,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             08-JAN-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             19-DEC-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 13-JUL-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Synthetic
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              AAB15224;
                                                                                                                                                                                                                                                                                              Sequence
                                                                                                                                                                                                                                                                                                                        Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   181
                                                                                                                                                                                                                                                                                                                                                         Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    RESULT 2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    AAB15224
a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                             a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
```

0

Gaps

;

120

9

```
The present sequence is a mutated version of a fusion protein created using the Hepatitis C virus (HCV) NS3 and NS4A protease enzymes. These proteins are both essential for the replication of the virus, acting to cleave its replicative proteins from the polyprotein produced from the HCV genome. Inhibitors of the two proteins should be effective as antiviral treatments of HCV Infection. This is useful as HCV can lead to chronic liver disease such as cirthosis, liver failure and liver cancer. The present invention concerns a number of NS3 mutants and NS3-NS4A well as enabling structural studies of the protease and proteins which can be used to identify inhibitors of this type, as well as enabling structural studies of the protease and protease: Inhibitor complexes. This sequence contains the alpha-helix0-7
                                                                                                                                                                                                                                                                                                                                                                       ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                             121 LVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVSTRGVAK 180
                                                                                                                                                                                                                                                                                                                                                                                                                                9
Modified hepatitis C virus (HCV) NS3 protease comprising at least 1 substitution of a hydrophobic alpha-helix 0 amino acid to a hydrophilic amino acid, useful for screening inhibitors that may treat hepatitis C
                                                                                                                                                                                                                                                                                                                                                                                                                  61 TSINGVLWTVYHGAGTRTIASPKGPVTQMYTNVDKDLVGWQAPQGSRSLTPCTCGSSDLY
                                                                                                                                                                                                                                                                                                                                                                                                  1 MKKKGSVVIVGRINLSGDTAYAQQTRGEQGTQKTSHTGRDKNQVEGEVQIVSTATQTFLA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Hepatitis; NS3 protease; viral replication; chronic liver disease;
liver failure; liver cancer; mutant; mutein.
                                                                                                                                                                                                                                                                                                                                                                       .:
O
                                                                                                                                                                                                                                                                                                                                        21; Length 197;
                                                                                                                                                                                                                                                                                                                                                                       Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Goldfarb V;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Hepatitis C virus NS4A-NS3 fusion protease #5.
                                                                                                                                                                                                                                                                                                                                     Score 1010; DB;
Pred. No. 9e-98;
0; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Zhang Y,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              $
                                                                              Claim 23; Fig 16; 66pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          AAB15223 standard; protein; 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   BRIM ) BRISTOL-MYERS SQUIBB CO
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    AVDFIPVESLETTMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    99.4%;
99.5%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Weinheimer S,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     99US-0115271
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        06-JAN-2000; 2000WO-US00345
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      19-DEC-2000 (first entry)
                                                                                                                                                                                                                                                                                                                                                                       Matches 196; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              WPI; 2000-465976/40.
N-PSDB; AAA73332.
                                                                                                                                                                                                                                                                                                                                                       Similarity
                                                                                                                                                                                                                                                                                                           197 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           WO200040707-A1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Wittekind M,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     38-JAN-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         13-JUL-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Synthetic.
                                                                                                                                                                                                                                                                                                             Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          AAB15223;
                                                                                                                                                                                                                                                                                                                                        Ouery Match
                                                                                                                                                                                                                                                                                variant.
                                                                                                                                                                                                                                                                                                                                                        Local
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               RESULT 3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               AAB15223
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ద
                                                                                                                                                                                                                                                                                                                                                                                                     õ
                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                               á
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ô
```

 $\sim$ 

```
Hepatitis; NS3 protease; viral replication; chronic liver disease; liver failure; liver cancer; mutant; mutein.
                                                                                                                                                           97.9%; Score 995; DB 21; 98.0%; Pred. No. 3.4e-96;
                                                                                                                                                                                                                                                                                                                                                                Hepatitis C virus NS4A-NS3 fusion protease #4.
                                                                                                                                                                         2; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Weinheimer S, Zhang Y,
                                                                                                                                                                                                                                                                                                                        Z
                                         15; 66pp; English
                                                                                                                                                                                                                                                                                                                      AAB15222 standard; protein; 197
                                                                                                                                                                                                                                                                               AVDFIPVESLETTMRSP 197
                                                                                                                                                                                                                                                                       AVDFIPVESLETIMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (BRIM ) BRISTOL-MYERS SQUIBB
                                                                                                                                                                                                                                                                                                                                                                                                                                                  2000WO-DS00345
                                                                                                                                                                                                                                                                                                                                                                                                                                                                99US-0115271
                                                                                                                                                                                                                                                                                                                                                  (first entry)
                                                                                                                                                                         Matches 193; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        WPI; 2000-465976/40
                                                                                                                                                                   Local Similarity
                                                                                                                                              197 AA;
                                                                                                                                                                                                                                                                                                                                                                                                  Hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                      WO200040707-A1
                                        23; Fig
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Wittekind M,
                                                                                                                                                                                                                                                                                                                                                                                                                                                  06-JAN-2000;
                                                                                                                                                                                                                                                                                                                                                  19-DEC-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                08-JAN-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                   13-JUL-2000
                                                                                                                                                                                                                                                                                                                                                                                                         Synthetic.
                                                                                                                                                                                                                  61
                                                                                                                                                                                                                                                                                                                                    AAB15222;
                                                                                                                                                                                                                                                                                    181
                                                                                                                                              Sequence
                                                                                                                                                                                                                                                                       181
                                                                                                                                                            Query Match
                                         Claim
                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                 ç
                                                                                                                                                                                                                            a
                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                      ç
                                                                                                                                                                                                                                                                                  셤
```

```
The present sequence is a mutated version of a fusion protein created using the Hepatitis C virus (HCV) NS3 and NS4A protease enzymes. These proteins are both essential for the replication of the virus, acting to cleave its replicative proteins from the polyprotein produced from the HCV genome. Inhibitors of the two proteins should be effective as antiviral treatments of the two proteins should be effective as antiviral treatments of the two proteins. Jiver failure and liver cancer. The present invention concerns a number of NS3 mutants and NS3-NS4A fusion proteins which can be used to identify inhibitors of this type, as well as enabling structural studies of the protease and protease: Inhibitor complexes. This sequence contains the alpha-helix0-1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            LVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVSTRGVAR 180
                                       Modified hepatitis C virus (HCV) NS3 protease comprising at least 1 substitution of a hydrophobic alpha-helix 0 amino acid to a hydrophilic amino acid, useful for screening inhibitors that may treat hepatitis {\cal C}
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                TSINGVLWTVYHGAGTRTIASPKGPVTQMYTNVDKDLVGWQAPQGSRSLTPCTCGSSDLY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1 MKKKGSVVIVGRINLSGDTAYAQQTRGEQGTQKTSHTGRDKNQVEGEVQIVSTATQTFLA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Hepatitis; NS3 protease; viral replication; chronic liver disease;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Goldfarb V;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    96.5%; Score 980; DB 21;
96.4%; Pred. No. 1.3e-94;
11ve 2; Mismatches 5;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Hepatitis C virus NS4A-NS3 fusion protease #3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              liver fallure; liver cancer; mutant; mutein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Wittekind M, Weinheimer S, Zhang Y,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     AAB15221 standard; protein; 197 AA
                                                                                                                                                Claim 23; Fig 14; 66pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               181 AVDFIPVESLETTMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          (BRIM ) BRISTOL-MYERS SQUIBB
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        06-JAN-2000; 2000WO-US00345.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 99US-0115271
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Best_Local Similarity 96.4
Matches 190; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                             197 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Hepatitis C virus
N-PSDB; AAA73331
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     NO200040707-A1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     19-DEC-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 08-JAN-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              13-JUL-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Synthetic
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              AAB15221;
                                                                                                                                                                                                                                                                                                                                                                                                                                                               Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                     variant
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                RESULT 5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     õ
                                                                                                                                                             The present sequence is a mutated version of a fusion protein created using the Hepatitis C virus (HCV) NS3 and NS4A protease enzymes. These proteins are both essential for the replication of the virus, acting to cleave its replicative proteins from the polyprotein produced from the Y genome. Inhibitors of the two proteins should be effective as antiviral treatments of HCV infection. This is useful as HCV can lead to chronic liver disease such as cirrhosis, liver failure and liver cancer. The present invention concerns a number of NS3 mutants and NS3-NS4A fusion proteins which can be used to identify inhibitors of this type, as well as enabling structural studies of the protease and
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                TSINGVLWTVYHGAGTRTIASPKGPVTQMYTNVDKDLVGWQAPQGSRSLTPCTCGSSDLY 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          9
                 Modified hepatitis C virus (HCV) NS3 protease comprising at least 1 substitution of a hydrophobic alpha-helix 0 amino acid to a hydrophilic amino acid, useful for screening inhibitors that may treat hepatitis C
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         61 TSINGVL#TVYHGAGTRTIASPRGPVTQMYTNVDKDLVGWQAPQGSRSLTPCTCGSSOLY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1 MKKKGSVVIVGRINLSGDTAYAQQTRGEQGTQKTSHTGRDKNQVEGEVQIVSTATQTFLA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Goldfarb V;
```

Gaps

ö

```
WPI; 2000-465976/40.
                                                                                                                                                                                                                                                                                             wild-type sequence
                                                                                                                                                                                                                                                                                                                       197 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Hepatitis C virus
                         N-PSDB; AAA73335
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     WO200040707-A1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    08-JAN-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                13-JUL-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Synthetic
                                                                                                                                                                                                                                                                                                                        Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       AAB15220;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     RESULT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                             셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  d
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ä
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         a
                                                                                                                                                                                                                                                                                                                                                             ö
                                                                                                                           The present sequence is a mutated version of a fusion protein created using the Hepatitis C vitus (HCV) NS3 and NS4A procease enzymes. These proteins are both essential for the replication of the vitus, acting to cleave its replicative proteins from the polyprotein produced from the HCV genome. Inhibitors of the two proteins should be effective as antiviral treatments of the two proteins should be effective as chronic liver disease such as cirrhosis, liver failure and liver cancer. The present invention concerns a number of NS3 mutants and NS3-NS4A well as enabling structural studies of the protease and proteins which can be used to identify inhibitors of this type, as well as enabling structural studies of the protease and protease. This sequence contains the alpha-helix0-1
                                                                                                                                                                                                                                                                                                                                                                                                                                                           TSINGVLWTVYHGAGTRTIASPKGPVTQMYTNVDKDLVGWQAPQGSRSLTPCTCGSSDLY 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                121 LVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVSTRGVAK 180
                                                                                                                                                                                                                                                                                                                                                                                                     1 MKKKGSVVIVGRINLSGDTAYAQQTRGEGCQETSQTGRDKNQVEGEVQIVSTAAQTFLA 60
                                      Modified hepatitis C virus (HCV) NS3 protease comprising at least 1 substitution of a hydrophobic alpha-helix 0 amino acid to a hydrophilic amino acid, useful for screening inhibitors that may treat hepatitis C
                                                                                                                                                                                                                                                                                                                                                                                        1 MKKKGSVVIVGRINLSGDTAYAQQTRGEQGTQKTSHTGRDKNQVEGEVQIVSTATQTFLA
                                                                                                                                                                                                                                                                                                                                                               Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Hepatitis; NS3 protease; viral replication; chronic liver disease;
liver failure; liver cancer; mutant; mutein.
                                                                                                                                                                                                                                                                                                                                                             ;
                                                                                                                                                                                                                                                                                                                                  Score 963; DB 21; Length 197;
Pred. No. 7.8e-93;
2; Mismatches B; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Weinheimer S, Zhang Y, Goldfarb V;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Hepatitis C virus NS4A-NS3 fusion protease #8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Ä
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (BRIM ) BRISTOL-MYERS SQUIBB CO.
                                                                                                       Claim 23; Fig 13; 66pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 AAB15226 standard; protein; 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    AVDFIPVESLETTMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    AVDFIPVESLETIMRSP 197
                                                                                                                                                                                                                                                                                                                                  94.8%;
94.9%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             06-JAN-2000; 2000WO-US00345
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (first entry)
                                                                                                                                                                                                                                                                                                                                               Best Local Similarity 94.9
Matches 187; Conservative
                                                                                                                                                                                                                                                                                                         197 AA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Hepatitis C virus
            N-PSDB; AAA73330
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       WO200040707-A1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Wittekind M,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        08-JAN-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    19-DEC-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  13-JUL-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Synthetic.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          AAB15226;
                                                                                                                                                                                                                                                                                                           Sequence
                                                                                                                                                                                                                                                                                                                                  Query Match
Best Local
                                                                                                                                                                                                                                                                                variant
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      RESULT 6
```

g

ö

셤

ö

g ò q

ò

```
The present sequence is a mutated version of a fusion protein created using the Hepatitis C virus (HCV) NS3 and NS4A protease enzymes. These proteins are both essential for the replication of the virus, acting to cleave its replicative proteins from the polyprotein produced from the HCV genome. Inhibitors of the two proteins should be effective as antiviral treatments of HCV infection. This is useful as HCV can lead to chronic liver disease such as cirrhosis, liver failure and liver cancer. The present invention concerns a number of NS3 mutants and NS3-NS4A well as enabling structural studies of the protease and proteins which can be used to identify inhibitors of this type, as well as enabling structural studies of the protease and protease inhibitor complexes. This sequence contains the alpha-helix0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   09
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           9
Modified hepatitis C virus (HCV) NS3 protease comprising at least 1 substitution of a hydrophobic alpha-helix 0 amino acid to a hydrophilic amino acid, useful for screening inhibitors that may treat hepatitis C
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             61 TSINGVLWTVYHGAGTRIIASPKGPVTQMYINVDKDLVGWQAPQGSRSLTPCTCGSSDLY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1 MKKKGSVVIVGRINLSGDTAYAQQTRGEQGTQKTSHTGRDKNQVEGEVQIVSTATQTFLA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Hepatitis: NS3 protease; viral replication; chronic liver disease;
liver failure; liver cancer; mutant; mutein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      .;
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Score 936; DB 21; Length 197;
Pred. No. 5.4e-90;
0; Mismatches 12; Indels (
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Hepatitis C virus NS4A-NS3 fusion protease #2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Ą
                                                                                                                                                                              Example 5; Fig 18; 66pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ő
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            AAB15220 standard; protein; 195
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       (BRIM ) BRISTOL-MYERS SQUIBB CO.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   AVDFIPVESLETTMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        AVDFIPVESLETIMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Query Match
Best Local Similarity 93.9%;
Matches 185; Conservative C
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           06-JAN-2000; 2000WO-DS00345
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         19-DEC-2000 (first entry)
```

ö

```
Wittekind M,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 409928482-A2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    28-NOV-1997;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           07-SEP-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       24-NOV-1998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                28-JUL-1998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           10-JUN-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    complexes
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      AAY24943;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               59
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         AAY 24943
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     RESULT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ð
                                                                                                                                                                                                                                                The present sequence is a mutated version of a fusion protein created using the Hepatitis C virus (HCV) NS3 and NS4A protease enzymes. These proteins are both essential for the replication of the virus, acting to cleave its replicative proteins from the polyprotein produced from the HCV genome. Inhibitors of the two proteins should be effective as antiviral treatments of HCV infection. This is useful as HCV can lead to chronic liver disease such as cirrhosis, liver failure and liver cancer. The present invention concerns a number of NS3 mutants and NS3-NS4A fusion proteins which can be used to identify inhibitors of this type, as well as enabling structural studies of the protease and processes inhibitor complexes. This sequence contains the alpha-helix0-1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           LVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVSTRGVAK 180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              TSINGVLWTVYHGAGTRT1ASPKGPVTQMYTNVDKDLVGWQAPQGSRSLTPCTCGSSDLY 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              substitution of a hydrophobic alpha-helix 0 amino acid to a hydrophilic amino acid, useful for screening inhibitors that may treat hepatitis C
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1 MKKKGSVVIVGRINLSGDTAYAQQTRGEQGTQKTSHTGRDKNQVEGEVQIVSTATQTFLA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Hepatitis; NS3 protease; viral replication; chronic liver disease;
liver failure; liver cancer.
                                                                                                         Modified hepatitis C virus (HCV) NS3 protease comprising at least
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Score 929; DB 21; Length 195;
Pred. No. 2.9e-89;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               9; Indels
Zhang Y, Goldfarb V;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Hepatitis C virus NS4A-NS3 fusion protease #1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             3; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          AAB15212 standard: protein; 195 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  (BRIM ) BRISTOL-MYERS SQUIBB CO.
                                                                                                                                                                                                             Claim 23; Fig 12; 66pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            AVDFIPVESLETTMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   AVDFIPVESLETIMRSP 195
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   91.4%;
Weinheimer S,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              06-JAN-2000; 2000WO-US00345
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        99US-0115271
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         19-DEC-2000 (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Matches 183; Conservative
                                         WPI; 2000-465976/40.
N-PSDB; AAA73329.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            195 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Hepatitis C virus Synthetic.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           WO200040707-A1
Wittekind M,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        08-JAN-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                119
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    AAB15212;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               59
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 179
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    variant
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Sest Local
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   RESULT 8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      AAB1521
NAME OF THE PROPERTY OF THE PR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
```

```
LVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVSTRGVAK 180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         61 TSINGVLWTVYHGAGTRTIASPRGPVTQMYTNVDKDLVGWQAPQGSRSLTPCTCGSSDLY 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    28
                                                                                                                                                                                                                                                                                                                                                                                                      The present sequence is a fusion protein created using the Hepatitis C virus (HCV) NS3 and NS4A protease enzymes. These proteins are both essential for the replication of the virus, acting to cleave its replication of the virus, acting to cleave its replicative proteins from the polyprotein produced from the HCV genome. Inhibitors of the two proteins should be effective as antiviral treatments of HCV infection. This is useful as HCV can lead to chronic liver disease such as cirrhosis, liver failure and liver cancer. The present invention concerns a number of NCS mutents and NS3-MCS4 fusion proteins which can be used to identify inhibitors of this type, as well as enabling structural studies of the protease and protease:inhibitor
                                                                                                                                                                    Modified hepatitis C virus (HCV) NS3 protease comprising at least 1 substitution of a hydrophobic alpha-helix 0 amino acid to a hydrophilic amino acid, useful for screening inhibitors that may treat hepatitis C
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1 MKKKGSVVIVGRIVLNG--AYAQQIRGLLGCIITSLIGRDKNQVEGEVQIVSTAAQTFLA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1 MKKKGSVVIVGRINLSGDTAYAQOTRGEQGTQKTSHTGRDKNQVEGEVQIVSTATQTFLA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  HCV; hepatitis C virus; single chain recombinant complex; linker;
NS3 protease; NS4A cofactor; NS4A-NS3 complex; serine protease;
hydrophobic domain; covalent complex; detection; inhibitor.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Length 195;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               13; Indels
Goldfarb V;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       88.8%; Score 902; DB 21; 91.9%; Pred. No. 2e-86;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1; Mismatches
   Zhang Y,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             AAY24943 standard; Protein; 665 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  HCV NS4A-NS3 complex SEQ ID NO:14.
                                                                                                                                                                                                                                                                                                                                             Example 2; Fig 10; 66pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 AVDFIPVESLETTMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 179 AVDFIPVESLETIMRSP 195
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  98US-0094331
97US-0067315
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 98WO-DS24528
   Weinheimer S,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Matches 181; Conservative
                                                            WPI; 2000-465976/40.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       195 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Hepatitis C virus.
Synthetic.
                                                                                                    N-PSDB; AAA73328
```

```
VTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVSTRGVAKA 181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               NSAA-NS3 complex comprising a central hydrophobic domain of native HCV hydrophobic domain of native HCV hydrophobic domain of native HCV NSAA peptide is tethered by the linker to the amino terminus of the HCV NSAA peptide is tethered by the linker to the amino terminus of the HCV NS3 protease domain. The present sequence represents a specifically claimed example of the above complex. The covalent NSAA-NS3 complexes are useful for structural determination and determination of mode of binding of HCV inhibitors by NMR spectroscopy. They can also be used for detecting inhibitors of the protease activity, the helicase activity and the ATPase activity of NS3. The covalent NSAA-NS3 complexes are more soluble, stable and active than the non-covalent protease-peptide complexes previously available.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            22 GSVVIVGRIILSGSGSITAYSQQTRGLLGCKKTSLTGRDKNQVEGEVQVVSTATQSFLAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SINGVLWTVYHGAGTRTIASPKGPVTQMYTNVDKDLVGWQAPQGSRSLTPCTCGSSDLYL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   5 GSVVIVGRINLSGD---TAYAQQTRGEQGTQKTSHTGRDKNQVEGEVQIVSTATQTFLAT
                                                                                                                                                                                                              describes a covalent hepatitis C virus (HCV)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             HCV; hepatitis C virus; single chain recombinant complex; linker; NS3 protease; NS4a cofactor; NS4a-NS3 complex; serine protease; hydrophobic domain; covalent complex; detection; inhibitor.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Query Match 85.9%; Score 872.5; DB 20; Length 665; Best Local Similarity 84.7%; Pred. No. 1.4e-82; Aatches 166; Conservative 16; Mismatches 11; Indels 3;
                                                                                                                           New hepatitis C virus covalent complexes
                                                                                                                                                                 Claim 6; Page 100-102; 211pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            AAY24942 standard; Protein; 665 AA.
                                          Weber PC,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     HCV NS4A-NS3 complex SEQ ID NO:13
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            182 VDFIPVESLETIMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  98US-0094331.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         98WO-0S24528
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     202 VDFVPVESMETTMRSP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             07-SEP-1999 (first entry)
                                          Malcolm BA, Taremi SS,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (SCHE ) SCHERING CORP
                                                                                                                                                                                                              present invention
(SCHE ) SCHERING CORP
                                                                                   WPI; 1999-385385/32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       665 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    W09928482-A2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       24-NOV-1998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  28-JUL-1998;
28-NOV-1997;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               10-JUN-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Synthetic.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      62
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         122
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      AAY24942;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     RESULT 11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          AAY24942
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         62 SINGVLWTVYHGAGTRIIASPRGPVT@MYTNVDKDLVGWQAPQGSRSLTPCTCGSSDLYL 121
                                                                                                                                                                                                                            The present invention describes a covalent hepatitis C virus (HCV) NS4A-NS3 complex comprising a central hydrophobic domain of native HCV NS4 serine protease domain, where the hydrophobic domain of native HCV NS3 perine protease domain, where the hydrophobic domain of native HCV NS4A peptide is tethered by the linker to the amino terminus of the HCV NS3 protease domain. The present sequence represents a specifically claimed example of the above complex. The covalent NS4A-NS3 complexes are useful for structural determination and determination of mode of binding of HCV inhibitors by NMR spectroscopy. They can also be used for detecting inhibitors of the protease activity, the helicase activity and the ATPase activity of NS3. The covalent NS4A-NS3 complexes are more soluble, stable and active than the non-covalent protease-peptide complexes previously available.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  81
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             5 GSVVIVGRINLSGD---TAYAQQTRGEQGTQKTSHTGRDKNQVEGEVQIVSTATQTFLAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               HCV; hepatitis C virus; single chain recombinant complex; linker;
NS3 protease; NS4A cofactor; NS4A-NS3 complex; serine protease;
hydrophobic domain; covalent complex; detection; inhibitor.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            3;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             86.2%; Score 875.5; DB 20; Length 665; 85.2%; Pred. No. 6.8e-83; tive 15; Mismatches 11; Indels 3;
                                                                                                                                               New hepatitis C virus covalent complexes
                                                                                                                                                                                     Claim 6; Page 90-92; 211pp; English
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              AAY24947 standard; Protein; 665 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         HCV NS4A-NS3 complex SEQ ID NO:18.
                                                              Weber PC,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                VDFIPVESLETIMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         VDFVPVESMETTMRSP 217
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    98US-0094331.
97US-0067315.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           98WO-US24528
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Query Match
Best Local Similarity 85.2%
Matches 167; Conscrvative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (first entry)
                                                              Taremi SS,
                                                                                                       WPI; 1999-385385/32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           665 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      W09928482-A2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              07-SEP-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           24-NOV-1998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    28-JUL-1998;
28-NOV-1997;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                10-JUN-1999
                                                              Malcolm BA,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Synthetic.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        AAY24947;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Sequence
```

셤

g δ qq

ò

ð

121

61 81

3;

ż

Yao

Weber PC,

~

```
Taremi SS,
                                           WPI; 1999-385385/32.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                216 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  W09928482-A2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   24-NOV-1998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 28-JUL-1998;
28-NOV-1997;
  Malcolm BA,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               07-SEP-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   10-JUN-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Synthetic.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              AAY24946;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          62
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   RESULT 13
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            AAY24946
  PARAMANA MARAMANA MAR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    â
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               VTRHADVIPVRRRGDSRCSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVSTRGVAKA 181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      142 VTRHADVIPVRRRGDSRGSLLSPRPVSYLKGSSGGPLLCPSGHAVGIFRAAVCTRGVAKA 201
                                                                                                                                                                                                                                     NSAA NB3 complex comprising a central hydrophobic domain of native MSAA Peptide, a linker, and an HCV NS3 serine protease domain, where the hydrophobic domain of native HCV NS3 serine protease domain, where the hydrophobic domain of native HCV NS3 periods is tethered by the linker to the amino terminus of the HCV NS3 protease domain. The present sequence represents a specifically claimed example of the above complex. The covalent NSAA NS3 complexes are useful for structural determination and determination of mode of binding of HCV inhibitors by NMR spectroscopy. They can also be used for detecting inhibitors of the protease activity, the helicase activity and the ATPase activity of NS3. The covalent NSAA-NS3 complexes are more soluble, stable and active than the non-covalent protease-peptide complexes previously available.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   83
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               22 GSVVIVGRIILSGSGSITAYSOOTRGLLGCIKTSLTGRDKNOVEGEVOVVSTATOSFLAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                5 GSVVIVGRINLSGD---TAYAQOTRGEQGTOKTSHTGRDKNQVEGEVQIVSTAIOTFLAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SINGVLWTVYHGAGTRTIASPKGPVTQMYTNVDKDLVGWQAPQGSRSLTPCTCGSSDLYL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     3; Gaps
                                                                                                                                                                                                                     present invention describes a covalent hepatitis C virus (HCV)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               HCV; hepatitis C virus; single chain recombinant complex; linker; NS3 protease; NS4A cofactor; NS4A-NS3 complex; serine protease; hydrophobic domain; covalent complex; detection; inhibitor.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    DB 20; Length 665;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     12; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  85.2%; Score 871.5; DB 2.85.2%; Pred. No. 1.8e-82; iive 14; Mismatches 12
                      ż
                                                                                                                       New hepatitis C virus covalent complexes
                         Yao
                                                                                                                                                                  Claim 6; Page 88-90; 211pp; English
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     AAY17880 standard; Protein; 216 AA
                      Weber PC,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 HCV NS4A-NS3 complex SEQ ID NO:4.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             VDFIPVESLETTMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        98US-0094331.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          98WO-US24528
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 07-SEP-1999 (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Query Match 85.8%
Best Local Similarity 85.2%
Matches 167; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         VDFVPVESMETTMRSP
                      SS,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  (SCHE ) SCHERING CORP.
                      Taremi
                                                                       WPI; 1999-385385/32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Hepatitis C virus.
Synthetic.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         665 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        28-JUL-1998;
28-NOV-1997;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       W09928482-A2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       24-NOV-1998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       10-JUN-1999
                      Malcolm BA,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               122
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             182
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                AAY17880;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           202
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Ω
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
```

```
SINGVLWTVYHGAGTRTIASPRGPVTQMYTNVDKDLVGWQAPQGSRSLTPCTCGSSDLYL 121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                NSAA NSI complex comprising a central hydrophobic domain of native HCV NSAA Peptide, a linker, and an HCV NSI serine protease domain, where the hydrophobic domain of native HCV NSAA serine protease domain, where the hydrophobic domain of native HCV NSAA peptide is tethered by the linker to the amino terminus of the HCV NSI protease domain. The present sequence represents a specifically claimed example of the above domplexes are useful for structural determination and determination of mode of binding of HCV inhibitors by NMR spectroscopy. They can also be used for detecting inhibitors of the protease activity, the helicase activity and the Arpase activity of NSI. The covalent NSAA.NSI complexes are more soluble, stable and active than the non-covalent protease-peptide complexes previously available.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             83
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          22 GSVVIVGRIILSGSGSITAYSQQTRGLLGCKKTSLTGRDKNQVBGEVQVVSTATQSFLAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                5 GSVVIVGRINLSGD---TAYAQQTRGEQGTQKTSHTGRDKNQVEGEVQIVSTATQTFLAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Gaps
                                                                                             present invention describes a covalent hepatitis C virus (HCV)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      HCV; hepatitis C virus; single chain recombinant complex; linker;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              NS3 protease; NS4A cofactor; NS4A-NS3 complex; serine protease; hydrophobic domain; covalent complex; detection; inhibitor.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  3;
                                                                                                                                                                                                                                                                                                                                                                                                                                     85.5%; Score 868.5; DB 20; Length 216; 85.1%; Pred. No. 7.6e-83; Live 15; Mismatches 11; Indels 3;
New hepatitis C virus covalent complexes
                                             Claim 6; Page 76-77; 211pp; English
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Taremi SS, Weber PC,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     AAY24946 standard; Protein; 665 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           HCV NS4A-NS3 complex SEQ ID NO:17.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               182 VDFIPVESLETIMRS 196
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          980S-0094331
970S-0067315
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  98WO-0524528
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                          Best Local Similarity 85.19
Matches 166; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  (SCHE ) SCHERING CORP.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Malcolm BA,
```

œ

```
WPI; 1999-385385/32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             24-NOV-1998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        28-JUL-1998;
28-NOV-1997;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        W09928482-A2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 07-SEP-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      10-JUN-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Synthetic.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     AAY17884;
                                                                                                                                                                                                                                                                                                                                                                                                                  Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                            Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   RESULT 15
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          AAY17884
  q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            THE STATE OF THE S
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       đ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     62 SINGVLWTVYHGAGIRTIASPKGPVTQMYTNVDKDLVGWQAPQGSRSLTPCTCGSSDLYL 121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            S CVNGVCWTVYHGAGSKTLAGPKGPITQMYTNVDQDLVGWQAPPGARSLTPCTCGSSDLYL 141
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          122 VTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVSTRGVAKA 181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 NS4A MS3 complex comprising a central hydrophobic domain of native HCV MS4A peptide. a linker, and an HCV NS3 berinc protease domain, where the hydrophobic domain of native HCV NS4A peptide is tethered by the linker to the amino terminus of the HCV NS3 protease domain. The present sequence represents a specifically claimed example of the above complex. The covalent NS4A-NS3 complexes are useful for structural determination and determination of mode of binding of HCV inhibitors by NMR spectroscopy. They can also be used for detecting inhibitors of the protease activity, the helicase activity and the Arpase activity of NS3. The covalent NS4A-NS3 complexes are more soluble, stable and active than the non-covalent protease-peptide complexes previously available.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          81
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            5 GSVVIVGRINLSGD---TAYAQQTRGEQGTQKTSHTGRDKNQVEGEVQIVSTATQTFLAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Gaps
                                                                                                                                                       present invention describes a covalent hepatitis C virus (HCV)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      HCV; hepatitis C virus; single chain recombinant complex; linker; NS3 protease; NS4A cofactor; NS4A-NS3 complex; serine protease; hydrophobic domain; covalent complex; detection; inhibitor.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         3;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                DB 20; Length 665;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Query Match 85.5%; Score 868.5; DB 20; Length Best Local Similarity 84.7%; Pred. No. 3.7e-82; Matches 166; Conservative 15; Mismatches 12; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Yao N;
                                                                    New hepatitis C virus covalent complexes
                                                                                                           Claim 6; Page 97-99; 211pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 AAY24941 standard; Protein; 665 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Weber PC,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               HCV NS4A-NS3 complex SEQ ID NO:12
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   UDEVPVESMETTMRSP 217
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              VDFIPVESLETTMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    97US-0067315
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 98WO-US24528
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           98US-0094331
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   07-SEP-1999 (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Taremi SS,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               (SCHE ) SCHERING CORP
                    WPI; 1999-385385/32
                                                                                                                                                                                                                                                                                                                                                                                                                                       665 A.A.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Hepatitis C virus
Synthetic.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             W09928482-A2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               24-NOV-1998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             28-JUL-1998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    28-NOV-1997;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           10-JUN-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Malcolm BA,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     202
                                                                                                                                                                                                                                                                                                                                                                                                                                       Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            AAY2494]
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ö
οy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            å
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Š
```

```
62 SINGVLWTVYHGAGTRTIASPRGPVTQMYTNVDKDLVGWQAPQGSRSLTPCTCGSSDLYL 121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  82 CVNGVCWTVYHGAGSKTLAGPKGPITQMYTNVDQDLVGWQAPPGARSLTPCTCGSSDLYL 141
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          122 VTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVSTRGVAKA 181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  The present invention describes a covalent hepatitis C virus (HCV) NS4A-NS3 complex comprising a central hydrophobic domain of native HCV NS4A-PST earlier protease domain, where the hydrophobic domain of native HCV NS4A peptide is tethered by the linker to the amino terminus of the HCV NS4A peptide is tethered by the linker to the amino terminus of the HCV NS3 protease domain. The present sequence represents a specifically claimed example of the above complex. The covalent NS4A-NS3 complexes are useful for structural determination and determination of mode of binding of HCV inhibitors by NMR spectroscopy. They can also be used for detecting inhibitors of the protease activity, the helicase activity and the ATPase activity of NS3. The covalent NS4A-NS3 complexes are more soluble, stable and active than the non-covalent protease-peptide complexes previously available.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              83
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  22 GSVVIVGRIILSGSGSITAYSQQTRGLIGCKITSLIGRDKNQVEGEVQVVSTATQSFLAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          5 GSVVIVGRINLSGD---TAYAQQTRGEQGTQKTSHTGRDKNQVEGEVQIVSTATQTFLAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         HCV; hepatitis C virus; single chain recombinant complex; linker;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 NS3 protease; NS4A cofactor; NS4A-NS3 complex; serine protease; hydrophobic domain; covalent complex; detection; inhibitor.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  <u>..</u>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          DB 20; Length 665;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  85.4%; Score 867.5; DB 20;
84.7%; Pred. No. 4.7e-82;
tive 15; Mismatches 12;
New hepatitis C virus covalent complexes
                                                                       Page 85-87; 211pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Weber PC,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 AAY17884 standard; Protein; 216 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 HCV NS4A-NS3 complex SEQ ID NO:8.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             182 VDFIPVESLETIMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     100 VDEVPVESMETIMESP 217
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        980S-0094331
970S-0067315
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            98WO-0524528
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Local Similarity 84.78 ses 166; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Malcolm BA, Taremi SS,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            (SCHE ) SCHERING CORP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               WPI; 1999-385385/32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Hepatitis C virus.
```

62 SINGVLWTVYHGAGTRTIASPKGPVTQMYTNVDKDLVGWQAPQGSRSLTPCTCGSSDLYL 121 5 GSVVIVGRINLSGD---TAYAQQTRGEQGTQXTSHTGRDKNQVEGEVQIVSTATQTFLAT 61 22 GSVVIVGRIILSGSGSITAYSQQTRGLLGCKKTSLTGRDKNQVEGEVQVVSTATQSFLAT 81 3; Gaps 85.2%; Score 865.5; DB 20; Length 216; 84.6%; Pred. No. 1.6e-82; 1ive 16; Mismatches 11; Indels 3; Best Local Similarity 84.5% Matches 165; Conservative 182 Query Match ô g ŏ 요 ò g ⋧

Search completed: August 30, 2003, 19:12:25 Job time : 45.6227 secs

1||:||||:|||||| 202 VDFVPVESMETTMRS 216

ğ

```
GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.
```

OM protein - protein search, using sw model

; Search time 16.2134 Seconds August 30, 2003, 19:02:22 Run on:

(without alignments) 1168.492 Million cell updates/sec

Title: Perfect score:

US-09-965-594-22 1016 1 MKKKGSVVIVGRINLSGDTA......VAKAVDFIPVESLETTMRSP 197 Sequence:

BLOSUM62 Gapop 10.0 , Gapext 0.5 Scoring table:

283308 seqs, 96168682 residues Searched: Total number of hits satisfying chosen parameters:

length: 0 length: 2000000000 Minimum DB seq Maximum DB seq

Post-processing: Minimum Match 0% Maximum Match 100% Listing first 45 summaries

PIR\_76:\*
1: pir1:\*
2: pir2:\*
3: pir3:\*
4: pir4:\* Database

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

major capsid prote	hypothetical prote	hypothetical prote	hypothetical prote	regulatory protein	autolysin, amidase	major acidic nucle	proteinase (EC 3.4	mosA protein - Rhi	hypothetical prote	hypothetical prote	hypothetical prote	transferrin-bindin	qenome polyprotein	genome polyprotein	genome polyprotein
VHBPLL	S57895	G90654	G85505	E64745	A11769	JC4013	PRLJHD	B53308	A81942	T08613	S72606	JN0821	GNWVY	GNWVYP	GNWVNE
7	7	7	~	~	~	~	~	-	~	7	~	~	-	~	ч
301	361	452	452	452	770	1023	323	333	539	619	302	911	3411	3411	3414
7.6	7.6	7.6	7.6	7.6	7.6	7.6	7.5	7.5	7.5	7.5	7.5	7.5	7.5	7.5	7.5
77.5	77.5	77.5	77.5	77.5	77.5	77	76.5	76.5	76.5	76.5	97	16	92	76	92
30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45

## ALIGNMENTS

	GNWVC3
genome N;Conta proted	genome polyprotein - hepatitis C virus (strain HCV-1) N:Contains: capsid protein C; envelope protein M; hepacivirin (EC 3.4.21.98) (nonstr protein NS4s; nonstructural protein NS4b; nonstructural protein NS5 C:Species: hepatitis C virus
C; Date C; Acce	C:Date: 30-Sep-1992 #sequence_revision 30-Sep-1992 #text_change 19-Jan-2001 C:Accession: A39166; PQ0403; PQ0404
R;Choc Proc.	R;Choo, Q.L.; Richman, K.H.; Han, J.H.; Berger, K.; Lee, C.; Dong, C.; Gallegos, C.; Proc. Natl. Acad. Sci. U.S.A. 88, 2451-2455, 1991
A;Tit]	le: Genetic organization and diversity of the hepatitis C virus.
A; Acce	A)ACCESSION: A39166 5 3100, MOID: 711/2020, FMID: 1040/0/4
A; Resi	A;Molecule (ype: mknA A;Residues: 1-3011 <cho></cho>
A; Cros	A; Cross references: GB: M62321; NID: 9329873; PIDN: AAA45676.1; PID: 9329874
J. Gen	kichan, s.w.; McCMish, r.; Holmes, E.C.; DOW, B.; Peutherer, J.F.; Follett, E.; Yap, J. Gen. Virol. 73, 1131-1141, 1992
A,Tit]	A:Title: Analysis of a new hepatitis C virus type and its phylogenetic relationship
A; Acce	A) ACCESSION: PO0403
A; Mole	A; Molecule type: genomic RNA
A; Resi	ddues: 15/7-1633 CGHA>
A; Expe	A7.EXperimental source: isolates E.bl6
A; Acce	A;Accession: PQ0404
A; Stat	A.Status: preliminary
A: MOLE	A:Rosidues: 1577-1633 <ch2></ch2>
A; Expe	srimental source: isolates E-b17
C; Supe	erfamily: hepatitis C virus genome polyprotein
C; Keys	Vords: ATP; capsid protein; envelope protein; glycoprotein; hydrolase; nonstruc
F; 116-	F.1.I.O.F.Count. Capsid to tell resolutions of the F.116 of 191/Product: envelope protein M *status predicted <epm></epm>
F; 192.	389/Product: major envelope protein E *status predicted <mee></mee>
F;390.	-729/Product: nonstructural protein NSI #status predicted <nsi></nsi>
F; 1007	1909/Fiboduct: Nonstitutuiai piotein maz #status predicted <msz <="" td=""></msz>
F;1230	F;1230-1237/Region: nucleotide-binding motif A (P-loop)
F; 131.	2-131/Keglon: nucleotide-binding motif B
F:1616	orisis/Region: DEAH motil 5-1862/Product: nonstructural protein NS4a #status predicted <n4a></n4a>
F;1863	F;1863-2013/Product: nonstructural protein NS4b *status predicted <a4b></a4b>
F; 2014	F:2014-3011/Product: nonstructural protein NS5 #status predicted <ns5></ns5>
io Carra	100,2009, 808,000,040,040,040,440,440,440,440,000,040,000,000,000,040,420,42
Best	Similarity 82.4%; Pred. No. 4.1e-68;
77 88	marches 100; conservative 9; mismarches 10; indeis 9; caps 1;

= = = :

..

<del>...</del>

COMP

```
genome polyprotein - hepatitis C virus (strain H)
N;Contains: capsid protein C; envelope protein M; hepacivirin (EC 3.4.21.98) (nonstruprotain NS4s; nonstructural protein NS5
C;Species: hepatitis C virus
A;Note: host Homo sapiens (man)
C;Date: 31-Dec-1992 *sequence_revision 31-Dec-1992 *text_change 19-Jan-2001
C;Accession: A36814; A41546
R;Inchauspe, G;; Zebedee, S;; Lee, D.H.; Sugitani, M.; Nasoff, M.; Prince, A.M.
submitted to GenBank, July 1992
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             A.Note: neither anino acid nor nucleotide sequence is given
G.Superfamily: hepatitis C virus genome polyprotein
G.Superfamily: hepatitis C virus genome polyprotein
G.Superfamily: hepatitis C virus genome polyprotein; glycoprotein; hydrolase; nonstruct
F.11-115/Product: capsid protein; envelope protein predicted <CPC>
F.116-119/Product: envelope protein M *status predicted <NED>
F.116-119/Product: envelope protein M *status predicted <NED>
F.116-119/Product: nonstructural protein NSI *status predicted <NED>
F.730-129/Product: hepaclvirin *status predicted <NSI>
F.1107-1615/Product: hepaclvirin *status predicted <NSI>
F.1230-1237/Region: nucleotide-binding motif A (P-loop)
                                                                                                                                                                                                                                                                                                                                          A:Description: Genomic structure of the human prototype strain H of hepatitis C virus A:Reference number: A36814
                                                                                                                                                                                                                                                                                                                                                                                                                        A; Wolecule type: genomic RNA
A; Residues: 1-3011 CINC>
A; Cross-references: GB:MID:g329737; PIDN:AAA45534.1; PID:g329738
A; Inchauspe, G; Zebedee, S.; Lee, D.H.; Sugitani, M.; Nasoff, M.; Prince, A.M.
Proc. Natl. Acad. Sci. U.S.A. 88, 10292-10296, 1991
A; Title: Genomic structure of the human prototype strain H of hepatitis C virus:
A; Reference number: A41546; MUID:92052256; PMID:1658800
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    F;1316-1319/Region: DEXH motif
                                                                                                                                                                                                                                                                                                                                                                                                       A; Accession: A36814
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          N;Contains: capsid protein C; envelope protein M; hepacivirin (EC 3.4.21.98) (nonstructuprotein NS4s; nonstructural protein NS5 C; Date: hepatitis C virus C; Date: 19-May-2000 #sequence_revision 19-May-2000 #text_change 19-Jan-2001 C; Accession: S40770; PC1285 F; Namoroto, H submitted to the EMBL Data Library, March 1992 A; Reference number: S40770
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              A;Molecule type: genomic RNA
A;Residues: 1-3011 <OKA>
A;Residues: 1-3011 <OKA>
A;Cross-references: EMBL:D10749; NID:g221586; PIDN:BAA01582.1; PID:g221587
A;Cross-references: EMBL:D10749; NID:g221586; PIDN:BAA01582.1; PID:g221587
B;Okamoto, H.; Okada, S.; Sugiyama, Y.; Yotsumoto, S.; Tanaka, T.; Yoshizawa, H.; Tsuda, Jpn. J. Exp. Med. 60, 167-177, 1990
A;Title: The 5'-terminal sequence of the hepatitis C virus genome.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          .005 RRGREILLGPADGMVSKGWRLLAPITAYAQQTRGLLGCIITSLIGRDKNQVEGEVQIVST 1064
                                                    54 ATQTFLATSINGVLWTVYHGAGTRTIASPKGPVTQMYTNVDKDLVGWQAPQGSRSLTPCT 113
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         114 CGSSDLYLVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAV 173
                                                                                                                                                        CGSSDLYLVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAV 173
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  54 ATQTFLATSINGVLWIVYHGAGTRTIASPKGPVTQMYTNVDKDLVGWQAPQGSRSLTPCT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           3 KKGSVVIVGRIN-----LSGDIAYAQQIRGEQGTQKTSHIGRDKNQVEGEVQIVST
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     F:1616-1862/Product: nonstructural protein NS4a #status predicted <NV4A>F:1863-2013/Product: nonstructural protein NS4b #status predicted <NV4B>F:2014-3011/Product: nonstructural protein NS5 #status predicted <NNS5>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               83.0%; Score 843.5; DB 1; Length 3011; 82.4%; Pred. No. 5c-68;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         19;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         8; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      174 STRGVAKAVDFIPVESLETTMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   174 STRGVAKAVDFIPVESLETTMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Ouery Match
Best Local Similarity 82.4
Matches 168; Conservative
                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   à
                                                                                                                                                                                                                                                               ò
                                                    ò
                                                                                                      g
                                                                                                                                                           ò
```

```
F;1616-1862/Product: nonstructural protein NS4a *status predicted <N4A>
F;1863-2013/Product: nonstructural protein NS4b *status predicted <NNB>
F;2014-3011/Product: nonstructural protein NS5 *status predicted <NS5>
F;2014-3011/Product: nonstructural protein NS5 *status predicted <NS5>
F;196,209,234,305,325,417,423,430,448,476,532,540,556,576,623,645,1213,1255,2041,224C
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    genome polyprotein - hepatitis C virus (strain Taiwan)
N;Contains: capsid protein C; envelope protein M; hepacivirin (EC 3.4.21.98) (nonstru-
protein NS4a; nonstructural protein NS4b; nonstructural protein NS5
C;Species: hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                A;Note: host demo sapiens (man)
C;Date: 31-Dec-1992 *sequence_revision 31-Dec-1992 *text_change 19-Jan-2001
C;Accession: A40244
                                                                                                                                                                                                                                                                                                                                                                               3 KKGSVVIVGRIN-----LSGDIAYAQQIRGEQGIQKISHIGRDKNQVEGEVQIVSI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 54 ATQTFLATSINGVLWTVYHGAGTRIIASPKGPVTQMYTNVDKDLVGWQAPQGSRSLTPCT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     114 CGSSDLYLVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      R;Chen, P.J.; Lin, M.H.; Tai, K.F.; Liu, P.C.; Lin, C.J.; Chen, D.S.
Virology 188, 102-113, 1992
                                                                                                                                                                                       Length 3011;
                                                                                                                                                                                                                                                              Indels
                                                                                                                                                                                       DB 1;
                                                                                                                                                                               Query Match 82.5%; Score 838.5; DB J
Best Local Similarity 81.4%; Pred. No. 1.4e-67;
Matches 166; Conservative 10; Mismatches 19
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     174 STRGVAKAVDFIPVESLETTMRSP 197
```

```
80.5%;
                                                                                                                          76.5%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Matches 155; Conservative
                                                                                                                            Best Local Similarity 76.5
Matches 156; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Similarity
  F;1616-1862/Product:
F;1863-2013/Product:
F;2014-3010/Product:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           174
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Query Match
Best Local S
                                                                                                     Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   genome
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              셤
                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                   8
                                                                                                                                                                                                                                                                                                                                                                                           οχ
                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ô
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                genome polyprotein - hepatitis C virus (strain JT)
N'Contains: capsid protein C; envelope protein M; hepacivirin (EC 3.4.21.98) (nonstructu protein NS4s: nonstructural protein NS4s: nonstructural protein NS5
C;Species: hepatitis C virus
C;Species: hepatitis C virus
C;Accession: A45573
R;Tanaka, T; Kato, N; Nakagawa, M; Octsuyama, Y; Cho, M.J.; Nakazawa, T; Hijikata, Virus Res. 23, 39-53, 1992
A;Title: Molecular cloning of hepatitis C virus genome from a single Japanese carrier: 8
A;Accession: A45573
A;Accession: P45573
A;Accession: P45573
A;Actular Coning Of hepatitis C virus genome from a single Japanese carrier: 8
A;Accession: P45573
A;Accession: P45573
A;Accession: P45573
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     A. Molecule type: DNA
A. Residues: 1-3010 crans
A. Cross-references: GB:D11168; GB:D01171; NID:g221612; PIDN:BAA01943.1; PID:g221613
A. Cross-references: HCV-JT
A. Note: Sequence extracted from NCBI backbone (NCBIN:106206, NCBIP:106207)
C. Superfamily: hepatitis C virus genome polyprotein
F:10-115/Product: capsid protein C #status predicted <PMS>
F:10-191/Product: major envelope protein B #status predicted <NES>
F:10-1006/Product: nonstructural protein NSI #status predicted <NS2>
F:1007-1615/Product: nonstructural protein NSI #status predicted <NS2>
F:1007-1615/Product: hepacivirin #status predicted <NS3>
F:130-1237/Region: nucleotide-binding motif B
F:1312-1317/Region: nucleotide-binding motif B
F:1316-1319/Region: DEXH motif
                                        F;1616-1862/Product: nonstructural protein NS4a #status predicted <N4A>
F;1863-2013/Product: nonstructural protein NS4b #status predicted <N4B>
F;2014-3010/Product: nonstructural protein NS5 #status predicted <NS5>
F;205,209,233,234,250,305,325,417,423,430,448,532,540,556,576,623,645,1213,1255,2041,207
A,Title: The Taiwanese hepatitis C virus genome: sequence determination and mapping the A,Reference number: A40244; MUID:92230206; PMID:1314449
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               54 ATQIFLATSINGVLWIVYHGAGIRTIASPKGPVTQMYTNVDKDLVGWQAPQGSRSLTPCT 113
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 114 CGSSDLYLVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAV 173
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     3 KKGSVVIVGRIN-----LSGDTAYAQQTRGEQGTQKTSHTGRDKNQVEGEVQIVST
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       .
6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Length 3010;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         18; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ä
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Score 826.5; DB 1
Pred. No. 1.8e-66;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       18; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         174 STRGVAKAVDFIPVESLETTMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         81.3%;
77.9%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Best Local Similarity 77.9
Matches 159; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
```

```
A; Molecule type: genomic RNA
A; Residues: 1-3010 <TAK
A; Residues: 1-3010 <TAK
A; Cross-references: EMBL:M58335; NID:9329770; PIDN:AAA72945.1; PID:9329771
C; Superfamily: hepatitis C virus genome polyprotein
C; Superfamily: hepatitis C **status predicted <PC>
F; 2-115, Product: capsid protein C **status predicted <PC>
F; 116-191/Product: major envelope protein E **status predicted <PC>
F; 130-739/Product: nonstructural protein N3 **status predicted <NS1>
F; 1307-1237/Region: nucleotide-binding motif A (P-loop)
F; 1312-1337/Region: nucleotide-binding motif A (P-loop)
F; 1316-1319/Region: nucleotide-binding motif B
F; 1616-1862/Product: nonstructural protein NS4** **status predicted <N48>
F; 166-1862/Product: nonstructural protein NS4** **status predicted <N85>
F; 1007-1615/Product: nonstructural protein NS5** **status predicted <N85>
F; 1007-1008, Product: nonstructural protein NS5** **status predicted <N85>
F; 1007-1008, Product: nonstructural protein NS4** **status predicted <N85>
F; 1007-1008, Product: nonstructural protein NS4** **status predicted <N85>
F; 1007-1008, Product: nonstructural protein NS5** **status predicted <N85>
F; 1007-1008, Product: nonstructural protein NS5** **status predicted <N85>
F; 1007-1008, Product: nonstructural protein NS5** **status predicted <N85>
F; 1007-1008, Product: nonstructural protein NS5** **status predicted <N85>
F; 1007-1008, Product: nonstructural protein NS5** **status predicted <N85>
F; 1007-1008, Product: nonstructural protein NS5*** **status predicted <N85>
F; 1007-1008, Product: nonstructural protein NS5*** **status predicted <N85>
F; 1007-1008, Product: nonstructural protein NS5*** **status predicted <N85>
F; 1007-1008, Product: nonstructural protein NS5*** **status predicted <N85>
F; 1007-1008, Product: nonstructural Protein NS5**** **status predicted <N85>**status Product <N85>**status Product <N85>**st
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 C;Accession: A38465
R;Takamizawa, A.; Mori, C.; Fuke, I.; Manabe, S.; Murakami, S.; Fujita, J.; Onishi, J. Virol. 65, 1105-1113, 1991
A;Title: Structure and organization of the hepatitis C virus genome isolated from h A;Reference number: A38465; MUID:91140698; PMID:1847440
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     N:Contains: capsid protein C; envelope protein M; hepacivirin (EC 3.4.21.98) (nonst protein N84s; nonstructural protein N85 (C; Species: hepatitis C virus (C; Species: 31-Mar-1992 *sequence_revision 31-Mar-1992 *text_change 19-Jan-2001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1005 RRGKEILLGPADSLEGRGLRLLAPITAYSQQTRGLLGCIITSLTGRDKNQVEGEVQVVST 1064
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1065 ATQSFLATCVNGVCWTVYHGAGSKTLAAPKGPITQMYTNVDQDLVGWPKPPGARSLTPCT 1124
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  113
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             114 CGSSDLYLVTRHADVIPVRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAV 173
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 53
                                                                                                                                                                                                                                                                                                                                                                                                                                 53
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             54 ATQTFLATSINGVLWTVYHGAGTRTIASPKGPVTQMYTNVDKDLVGWQAPQGSRSLTPCT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  RRGREILLGPADSIEGOGWRLLAPITAYAQQTRGLLGCIVTSLTGRDKNQVEGEVQVVST
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             54 ATQTFLATSINGVLWTVYHGAGTRTIASPKGPVTQMYTNVDKDLVGWQAPQGSRSLTPCT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             114 CGSSDLYLVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            3 KKGSVVIVGRIN-----LSGDTAYAQQTRGEQGTQKTSHTGRDKNQVEGEVQIVST
                                                                                                                                                                                                                                                                                                                                                                                                                            3 KKGSVVIVGRIN-----LSGDTAYAQQTRGEQGTQKTSHTGRDKNQVEGEVQIVST
                                                                                                                                                                                                                                                                                                                           Gaps
                <N4A>
nonstructural protein NS4a *status predicted <N4A:
nonstructural protein NS4b *status predicted <N4B:
nonstructural protein NS5 *status predicted <NS5>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ŝ
                                                                                                                                                                                                                                                                                                                           6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                80.1%; Score 813.5; DB 1; Length 3010; 76.0%; Pred. No. 2.7e-65;
                                                                                                                                                                                                                Length 3010;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      19; Indels
                                                                                                                                                                                                                                                                                                                           Indels
                                                                                                                                                                                                           Score 817.5; DB 1;
Pred. No. 1.2e-65;
20; Mismatches 19;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 21; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SIRGVAKAVDFIPVESLETIMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          polyprotein - hepatitis C virus
```

ġ,

ð 윤

```
A; Molecule type: genomic RNA
A; Residues: 1-547, T',549-621, V',623-624,'S',626-652,'DL',655-761,'T',763-782 < HOW>
A; Residues: 1-547, T',549-621,'V',623-624,'S',626-652,'DL',655-761,'T',763-782 < HOW>
A; Cross references: ExBL:x61591
A; Note: this sequence is inconsistent with the nucleotide translation
A; Note: the authors translated the codon AGG for residue 43 as Pro, TGG for residue
as TTP, and TTC for residue 771 as Ser
A; Note: sequence extracted from NCBI backbone (NCBIN:121747, NCBIP:121748)
C; Superfamily: hepatitis C virus genome polyprotein
C; Keywords: ATP: dytoprotein; hydrolase: nucleotide binding; P-loop; polyprotein; C; Keywords: ATP: dytoprotein; hydrolase: nucleotide depoly
F; 16-191/Product: envelope protein E $status predicted < NS1>
F; 190-729/Product: major envelope protein RS; *status predicted < NS1>
F; 130-127/Region: nucleotide-binding motif A (P-loop)
F; 1312-1317/Region: nucleotide-binding motif B
F; 1316-1319/Region: nucleotide-binding motif B
F; 16-191/Product: nonstructural protein NS4* *status predicted < NAA>
F; 165-1862/Product: nonstructural protein NS4* *status predicted < NAA>
F; 165-1862/Product: nonstructural protein NS4* *status predicted < NAA>
F; 166-1862/Product: nonstructural protein NS4* *status predicted < NAA>
F; 166-1862/Product: nonstructural protein NS4* *status predicted < NAA>
F; 166-1862/Product: nonstructural protein NS4* *status predicted < NAA>
F; 166-1862/Product: nonstructural protein NS4* *status predicted < NS5>
F; 100-209, 234, 250, 305, 417, 423, 448, 532, 540, 556, 576, 623, 645/Binding site: carbohydrate
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 genome polyprotein - hepatitis C virus (isolate EUH1480)
N;Contains: capsid protein C; envelope protein M; hepacivirin (EC 3.4.21.98) (nonstriprotein N94s; nonstructural protein N94s; nonstructural protein N95 C; Species: hepatitis C virus
C;Species: hepatitis C virus
C;Date: 19-May-2000 #sequence_revision 19-May-2000 #text_change 19-Jan-2001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          R;Chamberlain, R.W.; Adams, N.J.; Taylor, L.A.; Simmonds, P.; Elliott, R.M.
Biochem. Biophys. Res. Commun. 236, 44-49, 1997
A;Title: The complete coding sequence of hepatitis C virus genotype 5a, the predomin:
A;Reference number: JC5620; MUID:19365593; PMID:9223423
                                                                                                                                                                                                                                                                                                                                                                              A; Molecule type: genomic RNA
A; Residues: 1-3010 <HON>
A; Cross-references: EMBL: X61596; NID: 959478; PIDN: CAA43793.1; PID: 959479
A; Experimental source: isolate JK1 from an individual
R; Honda, M.; Kaneko, S.; Unoura, M.; Kobayashi, K.; Murakami, S.
Arch. Virol. 128, 163-169, 1993
A; Title: Sequence analysis of putative structural regions of hepatitis C virus isola
A; Reference number: A48332; MUID: 93119270; PMID: 8380322
A; Variety: isolate JK1
C: Date: 19-May-2000 #sequence_revision 19-May-2000 #text_change 23-Mar-2001
C: Accession: 518030; 533570; A48332; 518029
C: Accession: 518030; 533570; A48332; S18029
R: Honda, M.; Kaneko, S.; Masashi, U.; Kobayashi, K.; Murakami, S.
submitted to the EMBL Data Library, September 1991
A; Description: A whole genome of hepatitis C virus cDNA was isolated from a single
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              114 CGSSDLYLVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAV 173
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          54 ATQTFLATSINGVLWTVYHGAGTRTIASPKGPVTQMYINVDKDLVGWQAPQGSRSLTPCT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1065 ATQSFLATCVNGVCWTVYHGAGSKTLAGPKGPINQMTTNVDQDLVGWQAPSGAASLTPCT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                3 KRGSVVIVGRIN-----LSGDTAYAQQTRGEQGTQKTSHTGRDKNQVEGEVQIVST
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      78.9%; Score 801.5; DB 1; Length 3010; Ilarity 75.5%; Pred. No. 3.3e-64; Conservative 20; Mismatches 21; Indels 9;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CTRGVAKAVDFIPVESMETTMRSP 1208
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             174 STRGVAKAVDFIPVESLETTMRSP 197
                                                                                                                                                                                                                                                                                                   A; Reference number: S18028
A; Accession: S18030
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Best Local Sim
Matches 154;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1185
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 RESULT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      JC5620
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ó
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    A. Accession: PS0086
A. Molecule type: genomic RNA
A. Molecule type: genomic C. Stoperanily: hepatitis C virus genome polyprotein
C. Superfamily: hepatitis C virus genome polyprotein
C. Superfamily: hepatitis C virus genome polyprotein
C. Superfamily: hepatitis C virus genome polyprotein
F. 2-115/Product: capsid protein C *status predicted CPC>
F. 116-191/Product: major envelope protein E *status predicted CNSI>
F. 192-789/Product: nonstructural protein NSI *status predicted CNSI>
F. 1907-1615/Product: nonstructural protein NSI *status predicted CNSI>
F. 1907-1615/Product: nonstructural protein NSI *status predicted CNSI>
F. 1310-1317/Region: nucleotide-binding motif B
F. 1312-1317/Region: nucleotide-binding motif B
F. 1312-1317/Region: nucleotide-binding motif B
                                                                                                                                                                                                                                                                                                                              genome polyprotein - hepatitis C virus (strain J)
N;Contains: capsid protein C; envelope protein M; major envelope protein E; nonstructural protein NsContains: capsid protein C; envelope protein Ns4b; nonstructural protein NS4b; nonstructural protein NS5b; nonstructural protein NS4b; nonstructural protein NS5c; nonstructural NS5c; nonstructural
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          genome polyprotein - hepatitis C virus (isolate JK1)
N:Contains: capsid protein C; envelope protein M; hepacivirin (EC 3.4.21.98) (nonstructu protein NS4a; nonstructural protein NS4b; nonstructural protein NS5 C; Species: hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      A; Molecule type: genomic RNA
A; Residuces: 1-3010 (KAT>
A; Cross-references: GB:D90208; NID:g221610; PIDN:BAA14233.1; PID:g221611
R; Kato, N.; Ohkoshi, S.; Shimotohno, K.
R; Coc. Jpn. Acad. 658, 219-223, 1989
A; Title: Japanese isolates of the non-A, non-B hepatitis viral genome show sequence variance number: PS0085
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 F;1316-1319/Region: DEXH motif
F;1616-1662/Product: nonstructural protein NS4a #status predicted <N4A>
F;1663-2013/Product: nonstructural protein NS4b #status predicted <N4B>
F;2014-3010/Product: nonstructural protein NS5 #status predicted <NS5>
F;196,209,234,250,305,325,417,423,430,448,532,556,576,623,645,1213,1255,2041,2077,2240,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1005 RRGKEILLGPADSFGEQCWRLLAPITAYSQQTRGLLGCIITSLTGRDRNQVDGEVQVLST 1064
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             113
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  114 CGSSDLYLVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAV 173
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               53
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               3 KKGSVVIVGRINLSGD-----TAYAQQTRGEQGTQKTSHTGRDKNQVEGEVQIVST
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              54 ATQIFLAISINGVLWIVYHGAGIRIIASPKGPVIQMYINVDKDLVGWQAPQGSRSLIPCI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          .,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               DB 1; Length 3010;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     80.1%; Score 813.5; DB 1;
75.0%; Pred. No. 2.7e-65;
iive 23; Mismatches 19;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CTRGVAKAVDFVPVESMETTMRSP 1208
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      174 STRGVAKAVDFIPVESLETTMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    153; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               tches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                RESULT
S18030
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Ω
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            õ
```

1064

9

113

```
F;2018-3033/Product: nonstructural protein NSS *status predicted <NO5>
F:196,209,234,305,325,417,423,430,448,477,534,542,558,578,627,649,1091,1217,1259,20
                                                                                                                                       66.8%;
69.3%;
                                                                                                                                                                                                                           Matches 124; Conservative
                                                                                                                                  Query Match
Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             139
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         RESULT 11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   GNWVJ8
                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               MCD1303

Genome polyprotein - hepatitis C virus (isolate HC-J6)

N; Contains: capsid protein C; envelope protein M; hepacivirin (EC 3.4.21.98) (nonstructured protein NS4s; nonstructural protein NS4b; nonstructural protein NS5

C; Species: hepatitis C virus

C; Species: hepatitis C virus

C; Accession: J01303

R; Okamoto, H; Okada S: Suglyama, Y: Kurai, K: Ilzuka, H; Machida, A.; Miyakawa, Y.
J. Gen. Virol. 72, 2697-2704, 1991

A; Title: Nucleoride sequence—revision 19-May-2000 *text_change 17-Nov-2000

C; Accession: J01303

R; Okamoto, H; Okada S: Suglyama, Y: Kurai, K: Ilzuka, H; Machida, A.; Miyakawa, Y.
J. Gen. Virol. 72, 2697-2704, 1991

A; Title: Nucleoride sequence of the genomic RNA of hepatitis C virus isolated from a hum A; Reference number: J01303

A; Molecule type: genomic RNA

A; Residues: 1-3033 coKA>

A; R
                                                                                                                                                              A Experimental source: genotype 5a, which predominates in South Africa A. Experimental source: genotype 5a, which predominates in South Africa A. Note: the translation of the nucleotide sequence is not complete in this paper C. Superfamily: hepatitis C virus genome polyprotein a. Note: the translation of the nucleotide binding; P-loop; polyprotein; C. Keywords: ATP; glycoprotein; hydrolase; nucleotide binding; P-loop; polyprotein; F.2-115/Product: capsid protein C *status predicted <EP. F.116-191/Product: envelope protein E *status predicted <MEE> F.384-408/Region: hypervariable *status predicted <NS1> F.384-408/Region: hypervariable *status predicted <NS2> F.731-1007/Product: nonstructural protein NS1 *status predicted <NS2> F.731-1007/Product: nonstructural protein NS2 *status predicted <NS2> F.1231-1238/Region: nucleotide-binding motif B F.1331-1326/Region: nucleotide-binding motif B F.1317-1320/Region: nucleotide-binding motif B F.1317-1320/Region: nucleotide-binding motif B F.1317-1320/Region: nucleotide-binding motif B F.1317-1320/Region: nucleotide-binding motif F.1617-1863/Product: nonstructural protein NS4 *status predicted <N4A> F.1864-2014/Product: nonstructural protein NS5 *status predicted <NAS> F.1864-2014/Product: nonstructural protein NS5 *status predicted <NAS> F.2210-2249/Region: 1nterferon sensitivity determining *status predicted
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            54 AIQTFLATSINGVLWTVYHGAGTRTIASPKGPVTQMYTNVDKDLVGWQAPQGSRSLTPCT 113
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    114 CGSSDLYLVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAV 173
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    23
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           3 KKGSVVIVGRIN-----LSGDTAYAQQTRGEQGTQKTSHTGRDKNQVEGEVQIVST
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              F)1320-1323/Region: DEXH motif
F)1620-1866/Product: nonstructural protein NS4a *status predicted <N4A>
F)1867-2017/Product: nonstructural protein NS4b *status predicted <N4B>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              73.2%; Score 743.5; DB 1; Length 3014; 68.6%; Pred. No. 6.4e-59;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           30; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       25; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1186 CTRGVAKALEFVPVENLETTMRSP 1209
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 174 STRGVAKAVDFIPVESLETTMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       140; Conservative
                                     A,Molecule type: mRNA
A,Regidues: 1-3014 <CHA>
A,Cross-references: GB:Y13184
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 셤
```

```
A; Residues: 2678-2754 CHA>
A; Residues: 2678-2754 CHA>
A; Cross-references: DBBJ.1010134
A; Experimental source: isolate E-bl2
A; Experimental source: isolate E-bl2
B; Commun. 181, 279-285, 1991
B; Cochem. Biophys. Res. Commun. 181, 279-285, 1991
A; Title: Distribution of plural HCV types in Japan.
A; Title: Distribution of plural HCV types in Japan.
A; Reference number: PQ0554; MUID:92068204; PMID:1720309
A; Reference number: PQ0559
A; Referenc
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Cispeciaes: hepatitis C virus
Virology 188, 331-341, 1992
A Title: Full-length sequence of a hepatitis C virus genome having poor homology to
A:Reference number: A40250; MUID:92230232; PMID:1314459
A:Reference number: A40250
A:Residues: 1-3033 ACKA
A:References: GB:D1098B; GB:D01221; NID:922160B; PIDN:BAA01761.1; PID:9221609
B:Chan, S:W.; McComish, F.; Holmes, E.C.; Dow, B.; Peutherer, J.F.; Follett, B.; Yap
J. Gen. Virol. 73, 1131-1141, 1992
A:Reference number: P00393; MUID:92208B71; PMID:1316939
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                genome polyprotein - hepatitis C virus (strain HC-J8)
N;Contains: capsid protein C; envelope protein M; hepacivirin (EC 3.4.21.98) (nonst
protein NS4a; nonstructural protein NS4b; nonstructural protein NS5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              F;1620-1866/Product: nonstructural protein NS4a #status predicted <N4A>
F;1867-2017/Product: nonstructural protein NS4b #status predicted <N4B>
F;2018-3013/Product: nonstructural protein NS4 *status predicted <NS5>
F;196,209,233,299,305,417,423,430,448,477,534,542,558,578,627,649,1091,1217,1259,20
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ö
                                                                                                             ö
                                                                                                                                                                                                                                                                       79 IASPKGPVIQMYINVDKDLVGWQAPQGSRSLTPCTCGSSDLYLVTRHADVIPVRRRGDSR 138
                                                                                                                                                                                                     19 TAYAQQTRGEQGTQKTSHTGRDKNQVEGEVQIVSTATQTFLATSINGVLWTVYHGAGTRT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 GSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVSTRGVAKAVDFIPVESLETTMRSP
                                                                                                             Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ö
                                                                                                        ô
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Best Local Similarity 69.8%; Pred. No. 1.1e-52; Matches 125; Conservative 24; Mismatches 30; Indels
                                                                                                        28; Indels
Score 679; DB 1;
Pred. No. 4.7e-53;
                                                                                             27; Mismatches
```

Oy 53 TATQTFLATSINGVLWTVYHGAGTRTIASPKGP	RESULT 14 BB1104 nitrate/nitrite sensor protein (EC 2.7.3) NWB C.Species: Neisseria meningitidis C.Species: Neisseria meningitidis C.Accession: BB1104 C.Accession: BB1104 R.Tettelin, H.: Saunders, N.J.: Heidelberg, J.: Hickey, E.K.: Haft, D.H.: Salzberg, S.L.: White ri. H.; Olin, H.: Vamathevan, J.: Gill, J.: Scar Science 287, 1809-1815, 2000 A.Authors: Grandi, G.; Sun, L.: Smith, H.O.: Fr A.Title: Complete genome sequence of Neisseria A.Accession: BB1104 A.Accession: BB1104 A.Scatus: preliminary A:Nolecule type: DNA	A.Residues: 1-590 <tet> A.Cross-references: GB.AE002473, GB.AE002098, NA.Cross-references: GB.AE002473, GB.AE002098, NA.Cross-references: GB.AE002473, GB.AE002098, NA.Cross-reference: Serogroup B, strain MCS8 A.Gene: NMB1249 C.Genetics: NMB1249 C.Superfamily: nitrate/nitrite sensor protein nC.Keywords: autophosphorylation; phosphohistidif; F;395/Active site: His (phosphohistidine interm</tet>	Query Match 8.3%; Score 89.5; Best Local Similarity 20.8%; Pred. No. 2.9 Matches 45; Conservative 27; Mismatches	28 EGGTQKTSHTGRDKNQVEGEVQIVSTATQ 	0Y 88 OMTINUNDEDLYGWQAPQ	C.Date: US-May-2000 Sequence_revision US-May-2 C.Accession: C81911 C.Accession: C81911 F.Parkhill, J.; Achtman, M.; James, K.D.; Bentl F. Holroyd, S.; Jagels, K.; Leather, S.; Moule, Nature 404, 502-506, 2000 A.Title: Complete DNA sequence of a serogroup P. A.Reference number: A81775; MUID:20222556; PMII A;Status: preliminary
OY 19 TAYAOOTRGEOGTOKTSHTGRDKNOVEGEVOIVSTATGTFLATSINGVLMTVYHGAGTRT 78	PESULT 12  TOB841  POLYPOTCH - douroucouli hepatitis GB virus A  C; Species: douroucoull hepatitis GB virus A  C; Species: douroucoull hepatitis GB virus A  C; Date: 20-Sep-1999 *sequence_revision 20-Sep-1999 *text_change 17-Nov-2000  C; Accession: TOB841  R; Erker, JC: Desai, S.M.; Leary, T.P.; Chalmers, M.L.; Montes, C.C.; Mushahwar, I.K.  J, Gen. Virol. 79, 41-45, 1998  A; Title: Genomic analysis of two GB virus A variants isolated from captive monkeys.  A; Reference number: 216486; MUID:98120818; PMID:9460920  A; Accession: TOB841  A; Accession: TOB841  A; Molecule type: mRNA  A; Residues: 1-3005 <erk>  A; Molecule type: mRNA  A; Residues: 1-3005 <erk>  C; Superfamily: hepatitis C virus genome polyprotein  C; Keywords: polyprotein</erk></erk>	Query Match         24.7%:         Score 251;         DB 2;         Length 3005;           Best Local Similarity 34.1%:         Pred. NO. 4.2e-14;           Matches 56;         Conservative 29;         Mismatches 69;         Indels 10;         Gaps 3;           Qy 33 KTSHTGRDKNQVEGEVQIVSTATQTFLATSINGVLWITVYHGAGTRTIASPKGPVTQWTN 92	QY 93 VDKDLVGWQAPQGSRSLIPCTCGSSDLYLVTRHADVIPVRRRGDSRGSLLSPRPISYLKG 152 	Qy 153 SSGGPLLCPAGHAVGIRRAVSTRGVAKAVDFIPVES 189 	PESULT 13 T08839 Polyprotein - marmoset hepatitis GB virus A C; Species: marmoset hepatitis GB virus A C; Species: marmoset hepatitis GB virus A C; Date: 23-Jul-1999 #sequence_revision 23-Jul-1999 #text_change 17-Nov-2000 C; Accession: T08839 R; Erker, J.C.; Desal, S.M.; Leary, T.P.; Chalmers, M.L.; Montes, C.C.; Mushahwar, I.K. J. Gen. Virol. 79, 41-45, 1998 A; Title: Genomic analysis of two GB virus A variants isolated from captive monkeys. A; Recession: T08839 A; Status: translated from GB/EMBL/DDBJ A; Status: translated from GB/EMBL/DDBJ A; Molecule type: genomic RNA A; Residues: 1-2970 CERR> A; Cross-references: EMBL.FR023424; NID: 92828597; PIDN: AAC40501.1; PID: 92828598 C; Superfamily: hepatitis C virus genome polyprotein	Query Match  Query Match  Query Match  Best Local Similarity 27.8%; Score 246; DB 2; Length 2970;  Best Local Similarity 27.8%; Pred. No. 1.2e-13;  Matches 62; Conservative 39; Mismatches 80; Indels 42; Gaps 6;  Qy 3 KKGSVVIVGRINLSGDTAYAQOTRGEOGTOKTSHTGRDKNOVEGEVQIVS 52  ::     ::     ::       ::   Db 946 RRGDEVLIGVLNGVWELPPGFVPTAPVVVHHHGKGFFGVVKTSMTGWDETEHVGNVVUC 1005

```
narx
dine; phosphoprotein; phosphotransferas:
rmediate) *status predicted
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  tley, S.D.; Churcher, C.; Klee, S.R.; Mc, S.; Mungall, K.; Quail, M.A.; Rajandre
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     WA1418 [similarity] - Neisseria meningi
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   A strain of Neisseria menigitidis 2249:
ID:10761919
                                                                                                                                                                                                                                                                                                                .; Jeffries, A.C.; Nelson, K.E.; Eisen, te, O.; Fleischmann, R.D.; Dougherty, Barlato, V.; Masignani, V.; Pizza, M.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  NID:97226488; PIDN:AAF41629.1; PID:972
58
                                                                                                                                                                                                                                                                                                                                                                             Fraser, C.M.; Moxon, E.R.; Rappuoli, R.
a meningitidis serogroup B strain MC58.
ID:10710307
                                                                                                                                                                                                                                                   (B1249 [similarity] - Neisseria meningi
9
                                                                                     KGSSGGPLLCPAGHAVGIFRAAVSTR--- 176
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      | | | : | | CDEDDRILLQTLGRQLGVSLAGARQEEEK 378
                                                                    ------PRPISYLKGSSGGPLLCP 161
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    LATSINGVLWTVYHGAGTRT1ASPKGPVT 87
                                                                                                                                                                                                                                                                                2000 #text_change 19-Jan-2001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 -2000 *text_change 02-Feb-2001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  DB 2; Length 590;
.9;
es 79; Indels 65; Gaps
                                                                                                                                                        |:: | | |
REAITHTEAPP 1154
                                                                                                                                 VESLETTMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      : |||
|OMLET 414
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ESLET 192
```

```
A:Molecule type: DNA
A;Residues: 1-590 <PAR>
A;Residues: 1-590 <PAR>
A;Residues: 1-590 <PAR>
A;Cross=references: GB:AL162755; GB:AL157959; NID:g7379742; PIDN:CAB84658.1; PID:g738007
A;Experimental source: serogroup A, strain 22491
C;Genetics:
C;Genetics: NAN4418
C;Superfamily: nitrate/nitrite sensor protein narX
C;Repwords: autophosphorylation; phosphohistidine; phosphoprotein; phosphotransferase; s
F;395/Active site: His (phosphohistidine intermediate) *status predicted
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             | | | | : | | | | ::| : : | : : | : : | | ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| 
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     -----VIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVSTR--- 176
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            319 DLGKYHEEIFPIEYQNETLGRLLLSFPNGISLDEDDRILLQTLGRQLGVSLAGAKQEEER 378
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 28 EQGTQKTSHTGRDKNQVEGEVQIVSTATQTFLATSINGVLWTVYHGAGTRTIASPKGPVT 87
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Query Match 8.8%; Score 89.5; DB 2; Length 590;
Best Local Similarity 20.8%; Pred. No. 2.9;
Matches 45; Conservative 27; Mismatches 79; Indels 65; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               379 RLLAVLQERNLIAQGLHDSIAQALTFLNLQVQMLET 414
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   177
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Οý
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
```

Search completed: August 30, 2003, 19:20:32 Job time : 17.2134 secs

```
InterPro; IPR001410; DEAD.
InterPro; IPR002522; HCV_caps1d.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             EMBL; M62321; AAA45676.1; -. PIR; A39166; GNWVC3. PDB; 1A1V; 16-FEB-99. PDB; 1HEI; 25-NOV-98.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         MEROPS; S29.001; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            POLG_HCV1
ID POLG_HCV1
AC P26664;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            DESTRUCTION OF STREET AND DESCRIPTION OF STR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      P26664 h genome po P27958 h genome po P2996 h genome po P29663 h genome po P26663 h genome po P26660 h genome po P26660 h genome po P26661 h genome po P2661 h genome po P3372 treponema p P23936 hepatitis e P33436 hepatitis e P33436 hepatitis e P23591 heraticis e P2656 lactoroccus P2556 lactoroccus P2556 lactoroccus P2556 lactoroccus P25931 escherichia P26670 rhizobium m Q49768 mycobacteri m P03314 y genome po P9901 y genome po P9901 y genome po P990276 mus musculu P59240 mus musculu Q90607 gellus gall P59240 mus musculu Q90607 gellus gall
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                homo sapien
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             scrofa
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         musculu
                                                                                                                                                                                August 30, 2003, 18:01:52 ; Search time 9.75674 Seconds (without alignments) 949.524 Million cell updates/sec
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.
                                                                                                                                                                                                                                                                                                                                                                         1 MKKKGSVVIVGRINLSGDTA......vAKAVDFIPVESLETIMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                sns
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         шns
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Q28142
P11033
GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Total number of hits satisfying chosen parameters:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         127863 seqs, 47026705 residues
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         MOSA_RHIME
ERA_MYCLE
TBII_NEIMB
POLG_YEFVI
POLG_YEFVZ
POLG_TBEVW
TRYI_CHICK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      POLG_HCVH
POLG_HCVTW
POLG_HCVTW
POLG_HCVJR
P
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                GRAK_MOUSE
MCP_BPF41
MLTD_ECOLI
VPRT_SMRVH
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             NPH4_MOUSE
WDR9_HUMAN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      NX1B_BOVIN
GRAD_MOUSE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries
                                                                                                                         protein search, using sw model
                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Gapop 10.0 , Gapext 0.5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   length: 0
length: 2000000000
                                                                                                                                                                                                                                                                                                US-09-965-594-22
1016
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                DB
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SwissProt_41:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Query
Match Length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      33011
33011
33010
33010
33010
33010
33010
33010
33010
33010
33010
33010
33010
33010
33010
33010
33010
33010
33010
33010
33010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
34010
                                                                                                                                                                                                                                                                                                                                                                                                                                      BLOSUM62
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             844.5
838.5
826.5
817.5
813.5
813.5
679
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Score
                                                                                                                                                                                                                                                                                                                Title:
Perfect score:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Minimum DB seq
                                                                                                                                                                                                                                                                                                                                                                                                                                   Scoring table:
                                                                                                                            protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Database :
                                                                                                                                                                                                                                                                                                                                                                            Sequence:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Searched:
                                                                                                                                                                                         Run on:
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its way non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@lsb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SEQUENCE FROM N.A.

MEDLINE-91172826; PubMed-1848704;
Choo Q.-L., Richman K.H., Han J.H., Berger K., Lee C., Dong C.,
Choo Q.-L., Richman K.H., Han J.H., Berger K., Lee C., Dong C.,
Choo Q.-L., Richman K.H., Hadhar-Selby A., Barr P.J., Weiner A.J.,
Bardley D.W., Kuo G., Houghton M.;
Tegnetic organization and diversity of the hepatitis C virus.*;
Tegnetic organization and diversity of the hepatitis C virus.*;
Thypothologic Conformation and diversity of the hepatitis C virus.*;
Thypothologic Suddesiring A Possible A NS2A, NS2A, AND NS4B ARE
HYDORHOBIC SUGGESTING A POSSIBLE PREMARNE PROVICTION.
NS3 AND NS5 MAY PLAY A ROLE IN THE VIRAL RNA REPLICATION.
NS3 AND NS5 MAY PLAY A ROLE IN THE VIRAL RNA REPLICATION.
NS3 AND NS5 MAY PLAY A ROLE IN THE VIRAL RNA REPLICATION.
DIFFECURSOR POLYPOTEIN, commonly with Asp or Glu in the P6
posttion, Cys or Thr in Pl and Ser or Ala in Pl'.
C.:-CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
gibbon ape
hepatitis e
homo sapien
                                                                                                                                                                                                                                                                                                                                       aquifex aeo
hepatitis e
hepatitis e
                                                                                                                                  t genome po
gallus gall
                                                                                                                                                                                                                                                                                               triticum ae
                                                                                                                                                                                                             ustilago ma
                                                                                                                                                                                                                                                             synechocyst
                                                                                                                                                                                                                                                                                                                                                                                                                                                                     oryctolagus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     01-AUG-1992 (Rel. 23, Created)
01-AUG-1992 (Rel. 23, Last sequence update)
15-SEP-2003 (Rel. 42, Last annotation update)
Genome polyprotein (Contains: Capsid protein C (Core protein) (P22);
Envelope glycoprotein E1 (GP32) (GP35); Envelope glycoprotein E2
(GP68) (GP70) (NS1); Protein P7: Nonstructural protein NS2 (P21)
(EC 3.4.22.-); Protease/helicase NS3 (P70) (Hepacivirin)
(EC 3.4.21.98); Nonstructural protein NS4A (P4); Nonstructural protein
NS4B (P27); Nonstructural protein NS5A (P56); Nonstructural protein
NS5B (P66) (P70) (RNA-directed RNA polymerase) (EC 2.7.7.48)].
Hepatitis C virus (isolate 1) (HCV).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS: PROFEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MRNA.
                                                                              Q9p253 P
Q01299 t
Q01299 t
Q01299 t
P54856 t
P64856 t
Q01903 s
Q04531 t
Q006210 P
P28863 c
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY S29
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ALIGNMENTS
                                                                                                                 POLG_TBEVH
TRY2_CHICK
DPM1_USTWA
SUBI_SYNY3
CS12_WHEAT
FLII_AQUAE
VST2_HEVRH
VST2_HEVRH
MM03_RABIT
POL_GALV
VST2_HEVME
VP18_HUMAN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        STANDARD;
11165

14050

14050

14050

14050

14050

14050

14050

14050

14050

14050

14050

14050

14050

14050

14050

14050

14050

14050

14050

14050

14050

14050

14050

14050

14050

14050

14050

14050

14050

14050

14050

14050

14050

14050

14050

14050

14050

14050

14050

14050

14050

14050

14050

14050

14050

14050

14050

14050

14050

14050

14050

14050

14050

14050

14050

14050

14050

14050

14050

14050

14050

14050

14050

14050

14050

14050

14050

14050

14050

14050

14050

14050

14050

14050

14050

14050

14050

14050

14050

14050

14050

14050

14050

14050

14050

14050

14050

14050

14050

14050

14050

14050

14050

14050

14050

14050

14050

14050

14050

14050

14050

14050

14050

14050

14050

14050

14050

14050

14050

14050

14050

14050

14050

14050

14050

14050

14050

14050

14050

14050

14050

14050

14050

14050

14050

14050

14050

14050

14050

14050

14050

14050

14050

14050

14050

14050

14050

14050

14050

14050

14050

14050

14050

14050

14050

14050

14050

14050

14050

14050

14050

14050

14050

14050

14050

14050

14050

14050

14050

14050

14050

14050

14050

14050

14050

14050

14050

14050

14050

14050

14050

14050

14050

14050

14050

14050

14050

14050

14050

14050

14050

14050

14050

14050

14050

14050

14050

14050

14050

14050

14050

14050

14050

14050

14050

14050

14050

14050

14050

14050

14050

14050

14050

14050

14050

14050

14050

14050

14050

14050

14050

14050

14050

14050

14050

14050

14050

14050

14050

14050

14050

14050

14050

14050

14050

14050

14050

14050

14050

14050

14050

14050

14050

14050

14050

14050

14050

14050

14050

14050

14050

14050

14050

14050

14050

14050

14050

14050

14050

14050

14050

14050

14050

14050

14050

14050

14050

14050

14050

14050

14050

14050

14050

14050

14050

14050

14050

14050

14050

14050

14050

14050

14050

14050

14050

14050

14050

14050

14050

14050

14050

14050

14050

14050

14050

14050

14050

14050

14050

14050

14050

14
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  MCBI_TaxID=11104;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (RNA)(N).
```

```
9
                                                                                                              9
                                                                                                                                            g
                                                                                                                                                                               ö
                                           ò
                                                                           q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CAPSID PROTEIN (POTEWTIAL).
MATRIX PROTEIN (POTEWTIAL).
MAJOR ENVELORE PROTEIN E (POTENTIAL).
NONSTRUCTURAL PROTEIN NS1/E2 (POTENTIAL).
PROTEASE/HELICASE NS3 (POTEWTIAL).
NONSTRUCTURAL PROTEIN NS4A (POTEWTIAL).
NONSTRUCTURAL PROTEIN NS4A (POTEWTIAL).
NONSTRUCTURAL PROTEIN NS4A (POTEWTIAL).
RNA-DIRECTED RNA POLYMERASE (POTEWTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                             Polyprotein; Glycoprotein; Transferase; RNA-directed RNA polymerase; Core protein; Coat protein; Envelope protein; Helicase; ATP-binding; Transmembrane; Nonstructural protein; Hydrolase; Serine protease;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            SIMILARITY).
SIMILARITY).
SIMILARITY).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                             REMOVED FROM CAPSID PROTEIN C BY THE CELLULAR AMINOPEPTIDASE.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            .) (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Score 844.5; DB 1; Length 3011; Pred. No. 3.4e-70;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               (POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               65F8C9447FCE5AF9 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CHARGE RELAY SYSTEM (BY SI
CHARGE RELAY SYSTEM (BY SI
CHARGE RELAY SYSTEM (BY SI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              (GLCNAC.
(GLCNAC.
(GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ATP (POTENTIAL).
DECH BOX.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           N-LINKED
N-L
                                                                                              LINKED
                                                                              HCV_NS4a.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 AA; 327197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    83.18;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                645
2041
2077
2240
2364
2789
                                             IPR002518; H
IPR004109; H
IPR000745; H
                                                                                            IPR001490; F
                                                                                                                                                                                                                                                                                                                                                                                                                                                               3D-structure.
INIT_MET
                                                               InterPro;
InterPro;
InterPro;
InterPro;
                               InterPro;
InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ACT_SITE
ACT_SITE
NP_BIND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            TRANSMEM
ACT_SITE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CHAIN
```

53

9; Gaps

18; Indels

Mismatches

82.48;

Similarity

Local

Conservative

168;

Matches

3 KKGSVVIVGRIN-----LSGDTAYAQQTRGEQCTQKTSHTGRDKNQVEGEVQIVST

```
54 ATQTFLATSINGVLWTVYHGAGTRTIASPKGPVTQMYTNVDKDLVGWQAPQGSRSLTPCT 113
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                01-AUG-1992 (Rel. 23, Created)
01-AUG-1992 (Rel. 23, Last sequence update)
15-SEP-2003 (Rel. 42, Last annotation update)
Genome polyprotein (Contains: Capsid protein C (Core protein) (P22);
Envelope glycoprotein El (GP32) (GP35); Envelope glycoprotein E2
(GP68) (GP70) (MSI); Protein P7; Nonstructural protein NS2 (P21)
(EC 3.4.99-); Protease/helicase NS3 (P70) (Hepacivirin)
(EC 3.4.21.98); Nonstructural protein NS4A (P4): Nonstructural protein
NS4B (P27); Nonstructural protein NS5A (P56); Nonstructural protein
NS5B (P66) (P70) (RNA-directed RNA polymerase) (EC 2.7.7.48)].
Hepatitis C virus (isolate H) (HCV).
Viruses; SSRNA positive-strand viruscs, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS) OF 1192-1657.
MEDLINE-9015431; PubMed-9493270;
Kim J.L., Morganstern K.A., Gliffith J.P., Dwyer M.D., Thomson J.A.,
Murcko M.A., Lin C., Caron P.R.;
"Hepatitis C virus NS3 RNA helicase domain with a bound
oligonocleotide: the crystal structure provides insights into the mode
structure 6:89-100(1998).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  X-FAY CRYSTALLOGRAPHY (2.1 ANGSTROMS) OF 1207-1657.
MEDLINE-97331322; PubMed-9187654;
Yao N., Hesson T., Cable M., Hoop 2., Kwong A.D., Le H.V., Weber P.C.;
"Structure of the hepatitis C virus RNA helicase domain.";
Nat. Struct. Blol. 4:463-467(1997).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            -1- FUNCTION: PROTESAE NS IS RESPONSIBLE FOR THE CLEAVAGE OF NS2-NS3-1- FUNCTION: PROTESAE NS3 IS RESPONSIBLE FOR THE CLEAVAGE OF NS3-NSAE, NS4-NSAE, NS4B, NSAE, 
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           -:- FUNCTION: NSSA SEEMS TO HAVE A TRANSCRIPTIONAL ACTIVATORY ROLE.
-:- FUNCTION: NSSB IS A RNA-DEPENDENT RNA POLITMERASE THAT PLAYS AN
ESSEMTIAL ROLE IN THE VIRUS REPLICATION.
-:- CATALYTIC ACTIVITY: Hydrolysis of four peptide bonds in the viral
precursor polyprotein, commonly with Asp or Glu in the P6
position, Cys or Thr in P1 and Ser or Ala in P1'.
-:- CATALYTIC ACTIVITY: N nucleoside triphosphate - N diphosphate +
                                                                                                                                                     SUBDIVIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY 1 LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS: E1 AND E2. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MRNA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    U
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         "Genomic structure of the human prototype strain H of hepatitis virus: comparison with American and Japanese isolates."; Proc. Natl. Acad. Sci. U.S.A. 88:10292-10296(1991).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        [1]
SEQUENCE FROM N.A.
MEDLINE-92052256; PubMed-1658800;
Inchauspe G., Zebedee S., Lee D.H.H., Sugitani M., Nasoff M.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              PRT; 3011 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                        CTRGVAKAVDFIPVENLETTMRSP 1208
                                                                                                                                                                                                                                                                                                                                                                                                     174 STRGVAKAVDFIPVESLETTMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Hepacivirus.
NCBI_TaxID=11108;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        POLG_HCVH
ID POLG_HCVH
```

```
(POTENTIAL)
                 (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Length 3011;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         327142 MM; 772CBB29CCD94753 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Score 838.5; DB 1;
Pred. No. 1.2e-69;
                                                       (GLCNAC.
(GLCNAC.
(GLCNAC.
(GLCNAC.
(GLCNAC.
(GLCNAC.
                                                                                                                                                                                                                                                                                                    (GLCNAC.
(GLCNAC.
(GLCNAC.
            N-LINKED (

                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           82.5%;
81.4%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Ä
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Ouery Match
Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                11507
11511
11511
11514
11514
11514
11514
11514
11622
11622
11632
11632
11632
11632
11632
11632
11632
11632
11632
            CARBOHYD
TURN
TURN
TURN
TURN
TURN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  HELIX
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        STRAND
STRAND
STRAND
STRAND
STRAND
TURN
STRAND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    STRAND
HELIX
TURN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 STRAND
                 This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation-the European Bioinformatics Institute. There are no restrictions on its mach by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entitles requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       FIGURE TOO THE TOO THE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SIMILARITY).
SIMILARITY).
SIMILARITY).
PTM: THE STRUCTURAL PROTEINS C, EI AND E2 ARE PRODUCED BY PROTEOLYTIC PROCESSING BY THE HOST SIGNAL PEPTIDASES. SIMILARITY: THE NS2 PROTEASE BELONGS TO PEPTIDASE FAMILY U39. SIMILARITY: THE NS3 PROTEASE BELONGS TO PEPTIDASE FAMILY S29.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ..) (POTENTIAL)...) (POTENTIAL)...) (POTENTIAL)...)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            REMOVED FROM CAPSID PROTEIN C
CELLULAR AMINOPEPTIDASE.
CAPSID PROTEIN C.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             NONSTRUCTURAL PROTEIN NS2.
PROTEASE/HELICASE NS3.
NONSTRUCTURAL PROTEIN NS48.
NONSTRUCTURAL PROTEIN NS48.
NONSTRUCTURAL PROTEIN NS58.
POTENTIAL.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CHARGE RELAY SYSTEM (BY S. CHARGE RELAY SYSTEM (BY S. CHARGE RELAY SYSTEM (BY S. ATP (POTENTIAL).

DECH BOX.

N-LINNED (GLCNAC...) (PO N-LINNED (GLCNAC....) (PO N-LINNED (GLCNAC...) (PO N-LINNED (GLCNAC....) (PO N-LINNED (GLCNAC....) (PO N-LINNED (GLCNAC....) (PO N-LINNED (GLCNAC....) (PO N-LINNED (GLCNAC.....) (PO N-LINNED (GLCNAC.........)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ENVELOPE GLYCOPROTEIN E1.
ENVELOPE GLYCOPROTEIN E2.
PROTEIN P7.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             InterPro; INTO 2523; HCV-capsid.
InterPro; IPR002523; HCV-capsid.
InterPro; IPR002521; HCV-core.
InterPro; IPR002531; HCV-NS1.
InterPro; IPR002519; HCV-NS1.
InterPro; IPR004109; HCV-NS3.
InterPro; IPR004109; HCV-NS4.
InterPro; IPR004109; HCV-NS4.
InterPro; IPR004109; HCV-NS4.
InterPro; IPR001490; HCV-NS4.
InterPro; IPR007166; HCV-NS5.
InterPro; IPR007095; RNA_pol_DS_PS.
InterPro; IPR007095; RNA_pol_DS_PS.
InterPro; IPR007094; RNA_pol_DS_PS.
InterPro; IPR007094; RNA_pol_PSvir.
Pfam; PF01543; HCV_core; 1.
Pfam; PF01543; HCV_NS1; 1.
Pfam; PF01001; HCV_NS2; 1.
Pfam; PF01001; HCV_NS3; 1.
Pfam; PF01001; HCV_NS4; 1.
Pfam; PF01001; HCV_NS4; 1.
Pfam; PF01001; HCV_NS4; 1.
Pfam; PF01001; HCV_NS4; 1.
Pfam; PF00271; helicase_C; 1.
Pfam; PF00271; helicase_C; 1.
Pfam; PF00271; helicase_C; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   EMBL; M67463; AAA45534.1; -. PIR; A36814; GNWYCH. PDB; 1HEI; 25.NOV-98. PDB; 1A17; 16-FEB-99. PDB; 1A1R; 17-JUN-98.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                PDB: IAIV; 16-FEB-99.
PDB: IAIR: 17-JUN-98.
MEROPS; S29-001; --
MEROPS; U39-001; --
TRANSFAC: T04155; --
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               IPR001410;
IPR002522;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             192
384
810
1052
11658
117
1258
1107
1165
11165
1230
1316
1316
1316
1316
1316
1316
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               InterPro; InterP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CCHAIN
CCCARBOHYD
CCARBOHYD
CCA
```

us-09-965-594-22.rsp

```
InterPro;
InterPro;
InterPro;
InterPro;
InterPro;
InterPro;
                                                                                                                                                                                                                 InterPro;
                                                                                                                                                                                                                                         InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ACT_SITE
ACT_SITE
ACT_SITE
ACT_SITE
NP_BIND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CHAIN
TRANSMEM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ProDom; I
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CHAIN
CHAIN
CHAIN
CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                 Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                               Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Pfam;
     1064
                                                                                                                                                                                                                                                             This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bloinformatics and the EMB. outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-slb.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                           54 ATQTFLATSINGVLWTVYHGAGTRTIASPKGPVTQMYTNVDKDLVGWQAPQGSRSLTPCT 113
                                                         53
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              01-APR-1993 (Rel. 25, Created)
01-APR-1993 (Rel. 25, Last sequence update)
01-APR-1993 (Rel. 25, Last sequence update)
01-APR-1993 (Rel. 42, Last amototation update)
Genome polyprotein (Contains Capsid protein C (Core protein) (P22);
Envelope qlycoprotein El (GP32) (GP35); Envelope glycoprotein E2
(GP68) (GP70) (NG1); Protein P7; Nonstructural protein NS2 (P21)
(EC 3.4.22.-); Protease/hellcase NS3 (P70) (Hepacivirin)
(EC 3.4.21.98); Nonstructural protein NS4A (P4); Nonstructural protein
NS4B (P27); Nonstructural protein NS5A (P56); Nonstructural protein
NS5B (P60) (RNA-directed RNA polymerase) (EC 2.7.7.48)].
Hepatitis C virus (isolate Taiwan) (HCV).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     MEDLINE-92230206; PubMed-1314449;

MEDLINE-92230206; PubMed-1314449;

Chen P.J., Lin M.H., Tal K.F., Liu P.C., Lin C.J., Chen D.S.;

Chen P.J., Lin M.H., Tal K.F., Liu P.C., Lin C.J., Chen D.S.;

The Talwanese hepatitis C virus genome: sequence determination and mapping the 5' termini of viral genomic and antigenomic RNA.";

Virology 188:102-113(1992).

-I- FUNCTION: THE SMALL PROTEINS NS2A, NS2B, NS2B, NS4B AND NS5 MAY PLAY A ROLE IN THE VIRAL RNA REPLICATION.

-I- CATALYTIC ACTIVITY: Bydrolysis of four peptide bonds in the viral precursor polyprotein, commonly with Asp or Glu in the P6 position, Cys or Thr in Pl and Ser or Ala in Pl'.

-I- CATALYTIC ACTIVITY: N nucleoside triphosphate - N diphosphate + CANALYTIC ACTIVITY: N nucleoside triphosphate - N diphosphate + CANALYTIC ACTIVITY: N nucleoside triphosphate - N diphosphate + CANALYTIC ACTIVITY: N nucleoside triphosphate - N diphosphate - N di
                                                                                        3 KKGSVVIVGRIN----LSGDTAYAQQTRGEQGTQKTSHTGRDKNQVEGEVQIVST
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A LIPOPROTEIN ENVELOPE. CONSISTS OF TWO PROTEINS: PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C. AND MRNA.
  9; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY S29.
     Indels
     19:
10; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     PRT; 3010 AA
                                                                                                                                                                                                                                                                                                                                                                                                 1185 CTRGVAKAVDFIPVENLETTMRSP 1208
                                                                                                                                                                                                                                                                                                                                                                  174 STRGVAKAVDFIPVESLETTMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             EMBL; M84754; -; NOT_ANNOTATED_CDS.
PIR; A40244; GNWVTW.
Matches 166; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1N64; 25-FEB-03
1NS3; 08-APR-98
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      MEROPS; S29.001; ... MEROPS; U39.001; ...
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           NCBI_TaxID=31645;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          {RNA}(N).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              RESULT 3
POLG_HCTWW
AC P29846; VTW
AC P29846; VTW
AC P29846; VTW
AC P29846; VTW
AC D1-APR-1993
DT 15-SEP-2003
DE GENOME POLYP
DE GENOME POLYP
DE (EC 34-22.-
CC VITUSES; SSR
OC HEPATITIS CO VITUSES; SSR
OC HEPATITIS CC CATALYTIC
CC NS3 AND
CC -1- CATALYTIC
CC NS3 AND
CC -1- CATALYTIC
CC PROFILION
CC -1- CATALYTIC
CC PROFIEN
CC -1- SIMILARIC
CC -1- SIMILARIC
CC TAIS SWIGS FE
CC CATALYTIC
CC C CATALYTIC
CC CATALY
                                                                                                     g
                                                                                                                                                                                                   q
                                                                                                                                                                                                                                                          à
                                                                                                                                                                                                                                                                                                              a
                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                          ò
```

InterPro; IPR001410; DEAD

```
;
                                                                                                                                                                                                                                                                                                                                                      REMOVED FROM CAESID PROTEIN C BY THE CELLULAR AMINOPEPTIDASE.
CORE PROTEIN (POTENTIAL).
MATRIX PROTEIN (POTENTIAL).
MAJOR ENVELOPE PROTEIN E (POTENTIAL).
NONSTRUCTURAL PROTEIN NS1/E2 (POTENTIAL).
PROTEASE/HELICASE NS3 (POTENTIAL).
PROTEASE/HELICASE NS3 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4A (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4B (POTENTIAL).
RNA-DIRECTED RNA POLYMERASE (POTENTIAL).
RNA-DIRECTED RNA POLYMERASE (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Gaps
                                                                                                                                                                                                                                                                                                    Polyprotein; Glycoprotein; Transferase; RNA-directed RNA polymerase; Core protein; Coat protein; Envelope protein; Helicase; ATP-binding; Transmembrane; Nonstructural protein; Hydrolase; Serine protease;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SIMILARITY).
SIMILARITY).
SIMILARITY).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (POTENTIAL).
(POTENTIAL).
(POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Length 3010;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               (POTENT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          327047 MW; AAD267D55CDFE215 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CHARGE RELAY SYSTEM (BY
CHARGE RELAY SYSTEM (BY
CHARGE RELAY SYSTEM (BY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                DB 1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Score 826.5; DB 1
Pred. No. 1.6e-68;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (GLCNAC.
(GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (GLCNAC.
(GLCNAC.
(GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          (GLCNAC.
(GLCNAC.
(GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ATP (POTENTIAL)
DECH BOX.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          18; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          N-LINKED
N-LINKED
N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     N-LINKED
N-LINKED
N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              N-LINKED
N-LINKED
N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     N-LINKED
N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  N-LINKED
                                                                                   InterPro; IPR007095; RNA_POL_DS_PS.
InterPro; IPR007094; RNA_POL_PSvir.
Pfam: PF01543; HCV_capsid; 1.
                                                                                                                                                                                     HCV_NS3; 1.

HCV_NS4s; 1.

HCV_NS4s; 1.

HCV_NS5s; 1.

hclicase_C; 1.

viral_RdRP; 1.
                                                         HCV_NS5a.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             81.3%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Query Match
Best Local Similarity 77.99
Matches 159; Conservative
                                                                                                                               HCV_core;
HCV_env; 1
HCV_NS1; 1
                                                                                                                                                                                                                                                                                        DEXDC;
IPR002518;
IPR004109;
IPR000745;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        2041 204
2077 207
2240 224
2529 252
2788 278
3010 AA;
                                                           IPR002868;
                                            IPR001490;
                                                                                                                                                                                                                                                                          PD186062;
                                                                                                                                                                                     Pfam; PF02907; H
Pfam; PF01006; H
Pfam; PF01001; H
                                                                                                                                                                                                                                                                                        SM00487;
                                                                                                                                                                                                                                                              Pfam; PF00998;
                                                                                                                                                                                                                                                                                                                                                3D-structure.
```

S

```
DECH BOX.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  N-LINKED
                                                          InterPro; IPR004109; HCV_NS3.
InterPro; IPR001409; HCV_NS3.
InterPro; IPR001409; HCV_NS3.
InterPro; IPR001409; HCV_NS4.
InterPro; IPR001409; HCV_NS54.
InterPro; IPR001206; HCV_NS54.
InterPro; IPR001206; HCV_RRP.
InterPro; IPR001004; RNA_POL_DS_PS.
InterPro; IPR001004; RNA_POL_DS_PS.
InterPro; IPR001004; RNA_POL_PSVIT.
Pfam; PP01540; HCV_Carsid; I.
Pfam; PP01540; HCV_Carsid; I.
Pfam; PP01560; HCV_NS2; I.
Pfam; PP01006; HCV_NS3; I.
Pfam; PP010101; HCV_NS4; I.
Pfam; PP010101; HCV_NS4; I.
Pfam; PP010106; HCV_NS4; I.
Pfam; PP010106; HCV_NS4; I.
Pfam; PP010106; HCV_NS4; I.
Pfam; PP01010; HCV_NS4; I.
Pfam; PP01010; HCV_NS4; I.
Pfam; PP01010; HCV_NS4; I.
Pfam; PP01010; HCV_NS4; I.
Pfam; PP01011; HCV_NS4; I.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Best Local Similarity 76.5
Matches 156; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    2529
                                                                                                                                                                                                                                                                                                                                                          3D-structure,
                                                   InterPro;
                                       InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ACT_SITE
NP_BIND
SITE
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CHAIN
TRANSMEM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ACT_SITE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ACT_SITE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                       INIT_MET
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                         CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CHAIN
  the European Bioinformatics institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       This SWISS-PROI entry is copyright. It is produced through a collaboration
                 1125 CGSSDLYLVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPSGHVVGIFRAAV 1184
                                                               ATQTFLATSINGVLWTVYHGAGTRTIASPKGPVTQMYTNVDKDLVGWQAPQGSRSLTPCT 113
                                                                                                  CGSSDLYLVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAV 173
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     between the Swiss Institute of Bioinformatics and the EMBL outstation
3 KRGSVVIVGRIN-----LSGDTAYAQQTRGEQGTQKTSHTGRDKNQVEGEVQIVST 53
                                                                                                                                                                                                                                                              01-APR-1993 (Rel. 25, Created)
01-APR-1993 (Rel. 25, Last sequence update)
15-SEP-2003 (Rel. 25, Last sequence update)
01-APR-1993 (Rel. 26, Last sequence update)
66-0000 (Contains: Capsid protein C (Core protein) (P22);
87-000 (MSI); Protein E1 (GP32) (GP35); Envelope glycoprotein E2 (GP68) (GP76) (MSI); Protein P7; Nonstructural protein NS2 (P21)
(EC 3.4.22.-); Protease/helicase NS3 (P70) (Hepacivirin)
(EC 3.4.22.98); Nonstructural protein NS4A (P4); Nonstructural protein NS4B (P27); Nonstructural protein NS5A (P56); Nonstructural protein NS5A (P66); CP30) (RNA-directed RNA Polymerase) (EC 2.7.7.48)].
MS5B (P66) (P70) (RNA-directed RNA Polymerase) (EC 2.7.7.48)].
MS5B (P65) (P70) (RNA-directed NA Polymerase) (EC 2.7.7.48)].
                                                                                                                                                                                                                                                                                                                                                                                                                                                                            (RNA)(N).
SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLECCAPSID COVERED BY A SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLECCAPSID OF TWO PROTEINS:
PROTEIN M AND GLYCOPROTEIN E. THE NUCLECCAPSID IS A COMPLEX OF PROTEIN C AND MRNA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY S29.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    or send an email to license@isb-sib.ch).
                                                                                                                                                              1185 CTRGVAKAVDEVPESMETTMRSP 1208
                                                                                                                                                  174 STRGVAKAVDFIPVESLETTMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                       SEQUENCE FROM N.A. MEDLINE-92295714; Pubmed-1318627;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            EMBL; D11168; BAA01943.1; -
                                                                                                                                                                                                                                          STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1A1Q; 25-MAR-98.
1JXP; 14-JAN-98.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            MEROPS; S29.001; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          PIR; A45573; A45573
                                                                                                                                                                                                                                                                                                                                                                                                                                 NCBI_TaxID-31642;
                                                                                                                                                                                                                                         POLG_HCVJT
Q00269;
                                                  54
                                                                                                  114
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            <u>-</u>
                       셤
                                                ò
                                                                        a
                                                                                             ò
                                                                                                                         d
                                                                                                                                                    ö
                                                                                                                                                                        셤
```

```
ä
                                                                                                       CELLULAR AMINOPEPTIDASE.
CAPSID PROTEIN C (POTENTIAL).
MARTIX PROTEIN (POTENTIAL).
MAJOR ENVELOPE PROTEIN E (POTENTIAL).
NONSTRUCTURAL PROTEIN NSI/E2 (POTENTIAL).
                                                                                                                                                                                                                NON-STRUCTURAL PROTEIN NS2 (POTENTIAL).
PROTEASE/HELICASE NS3 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS48 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS48 (POTENTIAL).
RNA-DIRECTED RNA POLYMERASE (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Gaps
Polyprotein; Glycoprotein; Transferase; RNA-directed RNA polymerase;
                     Core protein; Coat protein; Envelope protein; Helicase; ATP-binding;
Transmembrane; Nonstructural protein; Hydrolase; Serine protease;
                                                                                    REMOVED FROM CAPSID PROTEIN C BY THE
                                                                                                                                                                                                                                                                                                                                                                             (BY SIMILARITY).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (POTENTIAL)
(POTENTIAL)
(POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ( POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         80.5%; Score 817.5; DB 1; Length 3010; 76.5%; Pred. No. 1.1e-67;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              3010 AA; 326573 MW; 94A1C77435D642BB CRC64;
                                                                                                                                                                                                                                                                                                                                                   CHARGE RELAY SYSTEM (BY
CHARGE RELAY SYSTEM (BY
CHARGE RELAY SYSTEM (BY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                N-LINKED (GLCNAC. . .)
                                                                                                                                                                                                                                                                                                                                                                                                                                                             N-LINED GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                            ATP (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Pred. No. 1.1e
20; Mismatches
```

53

3 KKGSVVIVGRIN-----LSGDTAYAQQTRGEQGTQKTSHTGRDKNQVEGEVQIVST

ò

InterPro; IPR001410; DEAD.

```
X *RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS) OF 1027-1210 AND 1678-1691.

X MEDLINE-98227846; PubMed-9568891;
Yan Y., Li Y., Munshi S., Sardana W., Cole J.L., Sardana M.,
Steinkuehler C., Tomei L., de Francesco R., Kuo L.C., Chen Z.;
Complex of NS3 protease and NS4A peptide of BK strain hepatitis C
T virus: a 2.2-A resolution structure in a hexagonal crystal form.;
Protein Sci. 7:837-847(1998).

C PROPOBIC SCI. 7:837-847(1998).

C PROPOBIC SUGGESTING A POSSIBLE MEBRANE-PELATED FUNCTION.
NS3 AND NS5 MAY PLAY A ROLE IN THE VIRAL RNA REPLICATION.
C CATALYTIC ACTIVITY: Hydrolysis of four peptide bonds in the viral precursor polyprotein, commonly with Asp or Glu in the P6
posttion, Cys or Thr in Pl and Ser or Ala in Pl'.

C CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
        1005 RRGREILLGPADSIEGQGWRLLAPITAYAQQTRGLLGCIVTSLTGRDKNOVEGEVQVVST 1064
                                                                    114 CGSSDLYLVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAV 173
                                                                                                                                                                                                     54 ATQTFLATSINGVLWTVYHGAGTRTIASPKGPVTQMYTNVDKDLVGWQAPQGSRSLTPCT
                                                                                                                                                                                                                                                                                                                                                          SEQUENCE FROM N.A.
MEDLINE-91140698; Pubmed-1847440;
Takamizawa A., Mori C., Puke I., Manabe S., Murakami S., Fujita J.,
Takamizawa A., Mori C., Puke I., Okayama H.;
"Structure and organization of the hepatitis C virus genome isolated
MEDLINE-97015088; PubMed-8861916;
Love R.A., Parge H.E., Wickersham J.A., Hostomsky Z., Habuka N.,
Moomaw E.W., Adachi T., Hostomska Z.;
Hip crystal structure of hepatitis C virus NS3 proteinase reveals
trypsin-like fold and a structural zinc binding site.";
Cell 87:331-342(1996).
                                                                                                                                                                                                                                                                                                                                                                                                                                                   MEDLINE-96235224; PubMed-8647104;
Borowski P., Heiland M., Oehlmann K., Becker B., Korneteky L.;
"Non-structural protein 3 of hepatitis C virus inhibits
phosphorylation mediated by cAMP-dependent protein kinase.";
Eur. J. Biochem. 237:611-618(1996).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 K-RAY CRYSTALLOGRAPHY (2.4 ANGSTROMS) OF 1027-1215.
PRT; 3010 AA.
                                                                                                                     1185 CTRGVAKAVDFIPVESMETTMRSP 1208
                                                                                                            174 STRGVAKAVDFIPVESLETTMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                              from human carriers.";
J. Virol. 65:1105-1113(1991).
                                                                                                                                                                                                                                                                                                                                                                                                                                            SEQUENCE OF 1487-1500
                                                                                                                                                                                                                                                                                                                                       NCBI_TaxID-11105;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  RNA](N)
                                                                                                                                                                                  POLG_HCVBK
P26663;
                                                                                                                                                                                                                                                                                                                               Hepacivirus
                                                                                                                                                                          POLG_HCVBK
                                                                                                                                                                RESULT 5
                                                                                                                                                                                   a
                                                                                                            ò
                                                                                                                              g
                                               g
                                                                    ò
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss institute of Bioinformatics and the EMBL outstation the European Bioinformatics institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CELLULAR AMINOPEPTIDASE.
CARSID PROTEIN ( POTENTIAL).
MATUR PROTEIN ( POTENTIAL).
MAJOR ENVELOPE PROTEIN E (POTENTIAL).
NONSTRUCTURAL PROTEIN NS1/E2 (POTENTIAL).
PROYEASE/HELICASE NS3 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4A (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4A (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4A (POTENTIAL).
RNA-DIRECTED RNA POLYMERASE (POTENTIAL).
BY A
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Polyprotein; Glycoprotein; Transferase; RNA-directed RNA polymerase; Core protein; Coat protein; Envelope protein; Helicase; ATP-binding; Transmembrane; Nonstructural protein; Hydrolase; Serine protease;
              LIPOPROFEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS: PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MRNA.
  SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    REMOVED FROM CAPSID PROTEIN C BY
                                                                          -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY S29.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CHARGE RELAY SYSTEM.
CHARGE RELAY SYSTEM.
CHARGE RELAY SYSTEM.
ATP (POTENTIAL).
DECH BOX.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           InterPro; IPR007095; RNA_pol_DS_PS.
InterPro; IPR007094; RNA_pol_DSvir.
Pfam; PF01543; HVV_capsid: 1.
Pfam; PF01542; HCV_core: 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   HCV_capsid.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   HCV_core.
HCV_NSI.
HCV_NSI.
HCV_NS2.
HCV_NS3.
HCV_NS4A.
HCV_NS4A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Pfam; PP01542; HCV_core; 1. Pfam; PP01539; HCV_core; 1. Pfam; PP01539; HCV_core; 1. Pfam; PP01569; HCV_NS2; 1. Pfam; PP01006; HCV_NS3; 1. Pfam; PP01006; HCV_NS4; 1. Pfam; PP01001; HCV_NS4a; 1. Pfam; PP01506; HCV_NS5a; 1. Pfam; PP019099; VITAL_KRRP; 1. Pr0Dom; PD186062; HCV_NS1; 1. SYART; SW00487; DEXDC; 11.
                                                                                                                                                                                                                                                                    EMBL; M58335; AAA72945.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     729
1006
1615
1615
1862
2013
3010
369
1107
11107
11165
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Interpro; IPR001410; D1
Interpro; IPR001522; HG
Interpro; IPR002521; HG
Interpro; IPR002519; HG
Interpro; IPR002519; HG
                                                                                                                                                                                                                                                                                      A38465; GNWYTC.
1A10; 25-MAR-98
1JXP; 14-JAN-98
1NS3; 08-APR-98
1C2P; 15-NOV-00
1C2P; 15-NOV-00
1CXF; 09-APR-02
1GX6; 10-APR-02
1GX6; 10-APR-02
1GX6; 10-APR-02
1GW4; 26-APR-99
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                IPR000745;
IPR001490;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             IPR002518;
IPR004109;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         IPR002868;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              IPR002166;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1616
1863
2014
347
1083
1107
1165
1230
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               3D-structure.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             InterPro;
InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       InterPro;
InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ACT_SITE
ACT_SITE
ACT_SITE
NP_BIND
SITE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       MEROPS; S
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           I RANSMEM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CHAIN
CHAIN
CHAIN
CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CHAIN
                                                                                                                                                                                                                                                                                                                                POB;
POB;
POB;
POB;
POB;
POB;
```

3010 AA

```
01-AUG-1992 (Rel. 23, Created)
01-AUG-1992 (Rel. 23, Last sequence update)
01-AUG-1992 (Rel. 21, Last sequence update)
01-AUG-1992 (Rel. 41, Last annotation update)
Genome polyprotein (Contains casial protein C (Core protein) (P22);
Envelope glycoprotein El (GP32) (GP35); Envelope glycoprotein E2
(GP68) (GP70) (NS1); Protein P7: Nonstructural protein NS2 (P21)
(EC 3.4.22.-); Protease/hellcase NS3 (P70) (Hepacivirin)
(EC 3.4.21.98); Nonstructural protein NS4A (P4); Nonstructural protein
NS4B (P27); Nonstructural protein NS5A (P56); Nonstructural protein
NS5B (P66) (P70) (RNA-directed RNA Polymerase) (EC 2.7.7.48)].
Hepatitis C virus (isolate Japanese) (HCV).
Viruses; SSRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                      STANDARD;
                                                                                                                                                                                                                                 Hepacivirus.
NCBI_TaxID=11116;
                                                   POLG_HCVJA
P26662;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     InterPro;
                             RESULT 6
                                                      DER NEUER HER DER NEUER DE PRESENTATION DE LA PRESENTATION DE LA PRESENTATION DE LA PRESENTATION DE LA PROPERCION DE LA PROPENCION DE LA PROPERCION DE LA PORTE DEL PORTE DEL PORTE DEL PORTE DEL PORTE DEL PORTE 
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         54 ATQTFLATSINGVLWTVYHGAGTRIIASPKGPVTQMYTNVDKDLVGWQAPQGSRSLJPCT 113
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CGSSDLYLVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAV 173
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             53
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          3 KKGSVVIVGRIN-----I,SGDTAYAQQTRGEQGTQKTSHTGRDKNQVEGEVQIVST
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Gaps
                                                              (POTENTIAL).
(POTENTIAL).
(POTENTIAL).
(POTENTIAL).
                                                                                                                 (POTENTIAL)
                                                                                                                                         (POTENTIAL)
(POTENTIAL)
                                                                                                                                                                              (POTENTIAL)
                                                    POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    6
                                       POTENTIAL)
                                                                                                                                                                                                        POTENTIAL)
                                                                                                                                                                                                                                POTENTIAL
                                                                                                                                                                                                                                          (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        80.1%; Score 813.5; DB 1; Length 3010; 76.0%; Pred. No. 2.7e-67;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  327189 MW; F8422D5ECCFDFD9C CRC64;
                                                                                                                (GLCNAC.
                                    (GLCNAC.
(GLCNAC.
(GLCNAC.
(GLCNAC.
(GLCNAC.
                                                                                                                                         (GLCNAC.
                                                                                                                                                                                                                                (GLCNAC.
                                                                                                                                                                              (GLCNAC.
                                                                                                                                                                   (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 21; Mismatches
N LINKED
                                                                                                                                                                                                                 N-LINKED
N-LINKED
N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               STRGVAKAVDFIPVESLETTMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Conservative
                                                                                                                                                                                                                                                                                                                                 645
2041
2077
2240
2529
2788
2788
1035
1047
1050
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1163
1163
1166
1174
1174
1188
1197
1202
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1688
                                                                                                                                                                                                                                                                                                                    940
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1189 120
1198 120
1203 120
1680 168
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Local Similarity
ses 155; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                             11101
11104
11120
11122
11129
11139
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1158
1162
1165
1168
                                                                                                                                                                                                                                                                                                                                            1082
1086
1090
1093
1095
                                                              CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
                                                                                                                CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1065
                                                                                                                                                                                          CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                114
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1125
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Query Match
                                                                                                                                                                                                                             CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SEQUENCE
                                       CARBOHYD
                                                   CARBOHYD
                                                                                                                                                                                                                   CARBOHYD
                                                                                                                                                                                                                                            CARBOHYD
                                                                                                                                                                                                                                                                                          STRAND
                                                                                                                                                                                                                                                                                                                                                                                             STRAND
STRAND
TURN
STRAND
STRAND
STRAND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    TURN
STRAND
STRAND
HELIX
TURN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              STRAND
TURN
STRAND
                                                                                                                                                                                                                                                                                                                   TURN
                                                                                                                                                                                                                                                      STRAND
                                                                                                                                                                                                                                                                               STRAND
                                                                                                                                                                                                                                                                                                                                                                      STRAND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 STRAND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      STRAND
                                                                                                                                                                                                                                                                                                                                             HELIX
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             HELIX
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Best Loca
Matches
                                                                                                                                                                                                                                                                                                                                                                                     TURN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    IURN
                                                                                                                                                                                                                                                                                                                                                            TURN
g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          õ
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bloinformatics institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   -- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY $29.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       -:- CATALYTIC ACTIVITY: Hydrolysis of four peptide bonds in the viral precursor polyprotein, commonly with Asp or Glu in the P6 position, Cys or Thr in P1 and Ser or Ala in P1'.
-:- CATALYTIC ACTIVITY: N nucleoside triphosphate - N diphosphate +
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      BY A
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY LIPOPROTEIN ENVELOPE CONSISTS OF TWO PROTEINS: PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MRNA.
                                                                                                                                                                                                                                                                                                                                                                            Obkoshi S., Shimotohno K.;

"Molecular structure of the Japanese hepatitis C viral genome.";

FEBS Lett. 280:325-328(1991).

1- FUNCTION: THE SMALL PROTEINS NS2A, NS2B, NS4A AND NS4B ARE

HYDROPHOBIC, SUGGESTING A POSSIBLE MEMBRANE-RELATED FUNCTION.

NS3 AND NS5 MAY PLAY A ROLE IN THE VIRAL RNA REPLICATION.
SEQUENCE FROM N.A.
MEDLINE-91088550; PubMed-2175903;
Kato N., Hijikata M., Ootsuyama Y., Nakagawa M., Ohkoshi S.,
Sugimura T., Shimotohno K.;
"Molecular cloning of the human hepatitis C virus genome from
Japanese patients with non-A, non-B hepatitis.;
Proc. Natl. Acad. Sci. U.S.A. 87:9524-9528(1990).
                                                                                                                                                                                                                                                                               DISCUSSION OF SEQUENCE.
MEDLINE-91192160; PubMed-1849488;
Kato N., Hijikata M., Nakagawa M., Ootsuyama Y., Muraiso K.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  HCV_capsid.
HCV_core.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  HCV_env.
HCV_NS1.
HCV_NS2.
HCV_NS3.
HCV_NS4a.
HCV_NS4b.
HCV_NS5a.
HCV_NS5a.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   EMBL; D90208; BAA14233.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  InterPro; IPR001410;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      InterPro; IPR002519;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   PIR, A39253; GNWVCJ.
HSSP; P26663; LJXP.
MEROPS; S29.001; -.
MEROPS; U39.001; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      InterPro; IPR002522;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  InterPro; IPR002521;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          InterPro; IPR002531;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          InterPro; IPR002518;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              InterPro; IPR004109;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  InterPro; IPR002166;
```

us-09-965-594-22.rsp

```
ATQTFLATSINGVLWTVYHGAGIRTIASPKGPVTQMYINVDKDLVGWQAPQGSRSLTPCT 113
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   114 CGSSDLYLVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAV 173
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       23
                                                                                                                                                          Probom: PD186062; HCV_NS1; 1.

SWART; SM00487; DEXDC; 1.

POLYPROTEIN: Glycoprotein: Transferase; RNA-directed RNA polymerase; Polyprotein: Glycoprotein: Envelope protein; Helicase; ATP-binding; Transmembrane; Nonstructural protein; Hydrolase; Setine protease.

INIT_MET 1 REMOVED FROM CAPSID PROTEIN C BY THE CELLULAR AMINOPEPITIOSE; CAPSID PROTEIN C (POTENTIAL).

CHAIN 116 191 MATRIX PROTEIN (POTENTIAL).

CHAIN 384 729 NON-STRUCTURAL PROTEIN NSI (POTENTIAL).

CHAIN 1006 NON-STRUCTURAL PROTEIN NSI (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      3 KKGSVVIVGRINLSGD-----TAYAQQTRGEQGTQKTSHTGRDKNQVEGEVQIVST
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               9; Gaps
                                                                                                                                                                                                                                                                                                    PROTEASE/HELICASE NS3 (POTENTIAL).
NONSTRUCTURAL PROTEIN NG4A (POTENTIAL).
NONSTRUCTURAL PROTEIN NG4B (POTENTIAL).
RNA-DIRECTED RNA POLYMERASE (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                   POTENTIAL.
CHARGE RELAY SYSTEM (BY SIMILARITY).
CHARGE RELAY SYSTEM (BY SIMILARITY).
CHARGE RELAY SYSTEM (BY SIMILARITY).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (POTENTIAL) (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Length 3010;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 MW; AA993794F46DB185
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        80.1%; Score 813.5; DB 1;
75.0%; Pred. No. 2.7e-67;
Live 23; Mismatches 19;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (GLCNAC.
(GLCNAC.
(GLCNAC.
(GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                 (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                       (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                        (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                               ATP (POTENTIAL)
DECH BOX.
                                                                                                                                                                                                                                                                                                                                                                                                                                N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         N-LINKED
N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                      N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  STRGVAKAVDFIPVESLETTMRSP 197
InterPro; IPR001650; Helicase_C.
InterPro; IPR007095; RNA_pol_DS_PS;
InterPro; IPR007094; RNA_pol_PSvir.
Pfam; PF01543; HCV_cap634; 1.
                                            Pfam; PP01542; HCV_COTE; 1. Pfam; PP01542; HCV_COTE; 1. Pfam; PP01542; HCV_COTE; 1. Pfam; PP01540; HCV_NS3; 1. Pfam; PP01006; HCV_NS3; 1. Pfam; PP010005; HCV_NS4; 1. Pfam; PP01001; HCV_NS4s; 1. Pfam; PP01506; HCV_NS4b; 1. Pfam; PP00971; helicase_C; 1. Pfam; PP00998; viral_RGRP; 1. ProDom; PD186652; HCV_NS1; 1. ProDom; PD186652; HCV_NS1; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   327017
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          2240
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Å.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Best Local Similarity
Matches 153; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          2240
2788
3010
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        645
2041
2077
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CARBOHYD
CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                    CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                           CARBOHYD
CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1125
                                                                                                                                                                                                                                                                                                                                                                                     ACT_SITE
NP_BIND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Query Match
                                                                                                                                                                                                                                                                                                                                                     FRANSMEN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CARBOHYD
                                                                                                                                                                                                                                                                                                                              CHAIN
                                                                                                                                                                                                                                                                                                                                          CHAIN
 윱
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation-the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        A SOUGHOUR PROMINE -9204440; PubMed-1658196;
A Okamoto H., Okada S.-I., Sugiyama Y., Kurai K., Lizuka H.,
A Machida A., Miyakawa Y., Mayumi M.;
Machida A., Miyakawa Y., Mayumi M.;
Mucleotide sequence of the genomic RNA of hepatitis C virus isolated
from a human carrier: comparison with reported isolates for conserved
and divergent regions.
T. J. Gen. Virol. 72:2697-2704 (1991).
C -!- FUNCTION: THE SMALL PROFISING NSZA, NSZB, NSAA AND NS4B ARE
HYDROPHOBIC. SUGGESTING A POSSIBLE MEMBRANE-RELATED FUNCTION.
NS3 AND NS5 AMY PLAY A ROLE IN THE VIRAL RNA REPLICATION.
NS3 AND NS5 AMY PLAY A ROLE IN THE VIRAL RNA REPLICATION.
OF CATALYTIC ACTIVITY: Hydrolysis of four peptide bonds in the viral prestucing componing with Asp or Glu in the P6
possition, Cys or Thr in Pl and Sar or Ala in Pl'.
C -I- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
                                                                                                                                                                                            01-AUG-1992 (Rel. 23, Created)
01-AUG-1992 (Rel. 23, Last sequence update)
28-FEB-2003 (Rel. 41, Last annotation update)
28-FEB-2003 (Rel. 41, Last annotation update)
Genome polyprotein [Contains: Capsid protein C (Core protein) (P22);
Envelope glycoprotein E1 (GP32) (GP35); Envelope glycoprotein E2 (GF68) (GP70) (NS1); Protein P7; Nonstructural protein NS2 (P21);
EC 3.4.22-7; Proteass/helicase NS3 (P70) (Hepacivirin)
EC 3.4.21-98); Nonstructural protein NSA (P56); Nonstructural protein NSAB (P27); Nonstructural protein NSAB (P56); Nonstructural protein NSAB (P56); Nonstructural protein NSAB (P56); Nonstructural protein NSAB (P56); Nonstructural protein NSAB (P50); Nonstructura
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             {RNA}(N):
-1-SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A LIPOPPOTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MRNA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            -! SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY S29.
                                                                                                                                                                       PRT; 3033 A.A.
1185 CTRGVAKAVDFIPVESMETTMRSP 1208
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            InterPro; IPR002166; HCV_RdRP.
InterPro; IPR001650; Hellcase_C.
InterPro; IPR007095; RNA_DOL_DS_PS.
InterPro; IPR007094; RNA_DOL_BS_PS.
Pfam; PF01543; HCV_caps1d; 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       HCV_capsid.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            HCV_NS4b,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               HCV_env.
HCV_NS1.
HCV_NS2.
HCV_NS3.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              HCV_NS4a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               EMBL; D00944; BAA00792.1; -.
                                                                                                                                                                       STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       IPR002522;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  IPR002521;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  IPR002519;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            InterPro; IPR001410;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             IPR002531;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     IPR002518;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      IPR004109;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              IPR000745;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               IPR001490;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          IPR002868;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       PIR; JQ1303; JQ1303.
HSSP; P27958; IHEI.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             NCBI_TaxID-11113;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   MEROPS; S29.001;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 MEROPS; U39.001;
                                                                                                                                                                    POLG_HCVJ6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          InterPro;
InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   InterPro;
InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       InterPro;
                                                                                                                                            POLG_HCVJ6
                                                                                                              RESULT 7
```

```
Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1094 LAGSRGPVTQMYSSAEGDLVGWPSPPGTKSLEPCTCGAVDLYLVTRNADVIPARRRGDKR 1153
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            79 IASPKGPVTQMYTNVDKDI,VGWQAPQGSRSLTPCTCGSSDLYLVTRHADVIPVRRRGDSR 138
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                78
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     GSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVSTRGVAKAVDFIPVESLETTMRSP 197
                                                                                                                                                                                                                                                                                                                       CAPSID PROTEIN C (POTENTIAL).
MATRIX PROTEIN (POTENTIAL).
MAJOR ENVELOPE PROFEIN E (POTENTIAL).
NONSTRUCTURAL PROTEIN NS1 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS2 (POTENTIAL).
PROTEASE/HELICASE NS3 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4A (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4B (POTENTIAL).
RNA-DIRECTED RNA POLYMERASE (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                19 TAYAQQTRGEQGTQKTSHTGRDKNQVEGEVQIVSTATQTFLATSINGVLWTVYHGAGTRT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              0; Caps
                                                                                                                                                                                                                            Polyprotein; Glycoprotein; Transferase; RNA-directed RNA polymerase; Core protein; Coat protein; Envelope protein; Helicase; ATP-binding; Transmembrane; Nonstructural protein; Hydrolase; Scine protease.
INIT_MET

| CELLULAR AMINOPEPTIDASE.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SIMILARITY).
SIMILARITY).
SIMILARITY).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       (POTENTIAL).
(POTENTIAL).
(POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        (POTENTIAL).
(POTENTIAL).
(POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          (POTENTIAL)
(POTENTIAL)
(POTENTIAL)
(POTENTIAL)
(POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       DB 1; Length 3033;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           28; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                POTENTIAL.
CHARGE RELAY SYSTEM (BY S
CHARGE RELAY SYSTEM (BY S
CHARGE RELAY SYSTEM (BY S
ATP (POTENTIAL).
DECH BOX.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 F957F5C1A273BE9E
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (GLCNAC.
(GLCNAC.
(GLCNAC.
(GLCNAC.
(GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               N-LINKED (GLCNAC.
N-LINKED (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ; Pred. No. 9e-55; 27; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Score 679;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 329165 MM;
                                                                                        PP01006; HCV_NS4a; 1.
PP01001; HCV_NS4b; 1.
PP01506; HCV_NS5a; 1.
PP00271; helicase_C; 1.
PP00998; viral_RdRP; 1.
PP0186062; HCV_NS1; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       66.8%;
69.3%;
                                    HCV_NS1; 1.
HCV_NS2; 1.
HCV_NS3; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Best Local Similarity 69.38
Matches 124; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    649
1091
2038
2811
                                                                                                                                                                                                                                                                                                                                                                                   733
010
619
619
866
2017
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  369
087
111
169
169
323
                                                                                                                                                                                                            SMART; SM00487; DEXDC;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Š
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1087
1111
1111
1169
1320
196
209
234
305
417
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               3033
                                                                                          Pfam; PF01006; Pfam; PF01001; Pfam; PF01506; P
                                                                          PF02907;
                                                                                                                                                                         Pfam; PF00998
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     139
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CARBOHYD
CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                TRANSMEM
ACT_SITE
ACT_SITE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ACT_SITE
NP_BIND
SITE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CARBOHYD
                                                                                                                                                                                           ProDom;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    RESULT 8
POLG_HCVJ8
                                                                                                                                                                                                                                                                                                                                                                                                                                                            CHAIN
                                                                                                                                                                                                                                                                                                                                                                                   CHAIN
CHAIN
CHAIN
CHAIN
                                                                                                                                                     Pfam;
                                                                                                                                                                                                                                                                                                                               CHAIN
                                                                                                                                                                                                                                                                                                                                                                   CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
DORR NO DORR N
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    a
```

3033 AA

STANDARD;

POLG\_HCVJ8 P26661;

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMEL CUISTATION. The European Bioinformatics institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
01-AUG-1992 (Rel. 23, Created)
01-AUG-1992 (Rel. 23, Last sequence update)
01-AUG-1992 (Rel. 21, Last sequence update)
08-FEB-2003 (Rel. 41, Last annotation update)
Genome polyprotein (Contains: Capsid protein C (Core protein) (P22);
Envelope glycoprotein E1 (GP32) (GP35); Envelope glycoprotein E2
(GP68) (GP70) (NS1); Protein P1 Northuctural protein NS2 (P21)
(EC 3.4.22.-); Protease/helicase NS3 (P70) (Hepacivirin)
(EC 3.4.22.-); Nonstructural protein NS4A (P4); Nonstructural protein
NS4B (P27); Nonstructural protein NS5A (P56); Nonstructural protein
NS5B (P66) (P70) (RNA-directed RNA polymerase) (EC 2.7.7.48)].
Hepatitis C virus (Isolate HG-J8) (HCV).
                                                                                                                                                                                                                              SEQUENCE FROM N.A.
MEDLINE-92230232; PubMed-1314459;
Okamoto H., Kurai K., Okada S.-I., Yamamoto K., Lizuka H., Tanaka T.,
Fukuda S., Tsuda F., Mishiro S.;
"Full-length sequence of a hepatitis C virus genome having poor
pomology to reported isolates: comparative study of four distinct
genotypes.":
                                                                                                                                                                                                                                                                                                                                                                                                                    the viral
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A LIPOPROTEIN ENTRLOPE. CHE ENVELOPE CONSISTS OF TWO PROTEINS: PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MENA.
                                                                                                                                                                                                                                                                                                                                                                  -1 - SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY S29.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Interpro; IPR002531; HCV_NS1.
Interpro; IPR002531; HCV_NS1.
Interpro; IPR00109; HCV_NS2.
Interpro; IPR001409; HCV_NS4.
Interpro; IPR001409; HCV_NS4b.
Interpro; IPR001406; HCV_NS5a.
Interpro; IPR002166; HCV_NS5a.
Interpro; IPR001098; RNA_POL_DS_PS.
Interpro; IPR001094; RNA_POL_PS_PS.
Interpro; IPR001094; RNA_POL_PS_PS.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       HCV_capsid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      HCV_core.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  InterPro; IPR002521; HCV_core
InterPro; IPR002519; HCV_env.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       EMBL; D10988; BAA01761.1; -. PIR; A40250; GNWVJ8.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Viral_RdRP;
                                                                                                                                                                                                                                                                                                                                                      Virology 188:331-341(1992)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              HCV_core;
HCV_env; 1
HCV_NS1; 1
HCV_NS2; 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            HCV_NS5a;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      InterPro; IPR001410;
InterPro; IPR002522;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       THEI.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   MEROPS; S29.001; -.
                                                                                                                                                                                              Hepacivirus.
NCBI_TaxID=11115;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            PF01506;
PF00998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              PF01539;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               PF01560;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               PF01001;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             140250; G
P27958;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    [RNA] (N).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          PIR; A40250
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Pfam;
Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Pfam;
```

```
.;
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             79 IASPKGPVTQMYTNVDKDLVGWQAPQGSRSLTPCTCGSSDLYLVTRHADVIPVRRRGUSR 138
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                19 TAYAQOTRGEQGTQKTSHTGRDKNQVEGEVQIVSTATQTFLATSINGVLWTVYHGAGTRT 78
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    139 GSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVSTRGVAKAVDFIPVESLETTHRSP 197
                                                                       CELLULAR ANIMORPATIONSE.
CAPSID PROTEIN ( POTENTIAL).
MAJOR ENVELOPE PROTEIN E (POTENTIAL).
NAJOR ENVELOPE PROTEIN E (POTENTIAL).
NONSTRUCTURAL PROTEIN NS1 (POTENTIAL).
PROTEASE/HELICASE NS3 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4A (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4B (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4B (POTENTIAL).
RNA DIRECTED RNA POLYMERASE (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Arabidopsis thaliana (Mouse-ear cress).
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         0; Gaps
     SHART; SHO487; DEDC; 1.
SHART; SHO487; DEDC; 1.
SHART; SHO487; DEDC; 1.
Core protein; Glycoprotein; Transferase; RNA-directed RNA polymerase;
Core protein; Coat protein; Envelope protein; Helicase; ATP-binding;
Transmembrane; Nonstructural protein; Hydrolase; Serine protease.
                                                                                                                                                                                                     POTENTIAL.
CHARGE RELAY SYSTEM (BY SIMILARITY).
CHARGE RELAY SYSTEM (BY SIMILARITY).
CHARGE RELAY SYSTEM (BY SIMILARITY).
                                                                                                                                                                                                                                                                                                                                                                                                                     (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                           (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                        (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                 (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                            (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (POTENTIAL)
                                                                                                                                                                                                                                                                                 (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                 (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                            (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                      (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Length 3033;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        30; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         N-LINKED (GLCNAC. . ) (POT W; 1A173E7E3381FD1A CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           HHOA_ARATH
052EL7; 049507;
10-CCT-2001 (Rel. 40, Created)
116-CCT-2001 (Rel. 40, Last sequence update)
116-CCT-2001 (Rel. 40, Last sequence update)
116-CCT-2001 (Rel. 40, Last annotation update)
116-CCT-2001 (Rel. 40, Last precursor (BC 3.4.21.-).
116-CCT-2001 (Rel. 40, CRIOTOPLAST PRECURSOR (BC 3.4.21.-).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (GLCNAC. . .)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              66.4%; Score 675; DB 1;
69.8%; Pred. No. 2.1e-54;
11ve 24; Mismatches 30
                                                                                                                                                                                                                                                                                                                     (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                       (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                    (GLCNAC
                                                                                                                                                                                                                                                                                 (GLCNAC
                                                                                                                                                                                                                                                                                                                                           (GLCNAC
                                                                                                                                                                                                                                                    ATP (POTENTIAL).
DECH BOX.
                                                                                                                                                                                                                                                                                                      (GLCNAC
                                                                                                                                                                                                                                                                                                                                                          (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                      (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                  (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                            (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                            (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                          (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  (GLCNAC
                                                                                                                                                                                                                                                                                                                               N-LINKED
N-LINKED
N-LINKED
                                                                                                                                                                                                                                                                                                                                                                   N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
                                                                                                                                                                                                                                                                                                        N-LINKED
N-LINKED
                                                                                                                                                                                                                                                                                  N-LINKED
                                                                                                                                                                                                                                                                                            N-LINKED
 ProDom; PD186062; HCV_NS1; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       330177
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Local Similarity 69.8% les 125; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        AA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CARBOHYD
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                            CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                         CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                     CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Query Match
                                                                                                                                                                                                                                                                                 CARBOHYD
                                                                                                                                                                                                                                                                                                         CARBOHYD
                                                                                                                                                                                                                                                                                                                       CARBOHYD
                                                                                                                                                                                                                                                                                                                                 CARBOHYD
                                                                                                                                                                                                                                                                                                                                              CARBOHYD
                                                                                                                                                                                                                                                                                                                                                           CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                      CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                   CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                 CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                            CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CARBOHYD
                                                                                                                                                                                                         TRANSMEM
                                                                                                                                                                                                                                                                                              CARBOHYD
                                                                                                                                                                                                                                                        NP_BIND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  RESULT 9
HHOA_ARATH
                                                                                                    CHAIN
CHAIN
CHAIN
CHAIN
CHAIN
                                                                                                                                                                               CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Aatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  οy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   õ
```

```
190 LET 192
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Best
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    11;
                                                                    entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
                                       Usage by and for commercial
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 151
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   KTSPSVVYIEAIELPKTSSGDILTDEENGKIEGTGSGFVWDKLGHI-----VTNYHVIA 148
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           52 STATQTFLATSINGVLWTVYHGAGTRTIASPKGPVTQMYTNVDKDLVGWQAPQGSRSLTP 111
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   202 VVLGTSNDLRVGQSCFAI----GNPYGYENTLTIGVVSGLGREIPSPNGKSISEAIQTD 256
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     2 KKKGSVVIVGRINL----SGDTAYAQQTRGEQGT-----QKTSHTGRDKNQVEGEVQIV 51
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SEQUENCE FROM N.A.
STRAIN-ATCC 15692 / PAO1;
STRAIN-ATCC 15692 / PAO1;
STRAIN-LOUGH STRAIN STRAIN STRAIN-ATCC 15692 / PAO1;
STOWN STRAIN-STRAIN STRAIN STRAI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 112 CTCGSSDLYLVTRHADVIPVRRRGDSRG------SLLSPRPISYLK----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    56; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                      Chloroplast: Thylakoid: Transit peptide. CHLOROPLAST (POTENTIAL).
non-profit institutions as long as its content is in
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   257 ADINSGNSGGPLLDSYGHTIGVNTAIFTRKGSGMSSGVNFAIPIDTVVRTV 307
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ----GSSGGPLLCPAGHAVGIFRAAVSTR--GVAKAVDF-IPVESLETIM 194
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       NATURE 406:959-964(2000).
-1- SIMILARITY: BELONGS TO THE POLYPRENYL P-HYDROXYBENZOATE / PHENYLACRYLIC ACID DECARBOXYLASES FAMILY.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CHARGE RELAY SYSTEM (POTENTIAL).
CHARGE RELAY SYSTEM (POTENTIAL).
CHARGE RELAY SYSTEM (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               DB 1; Length 321;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    89; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      R -> G (IN REF. 1).
68DB81E0BD27A7A7 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 16-OCT-2001 (Rel. 40, Last sequence update)
16-OCT-2001 (Rel. 40, Last annotation update)
Probable aromatic acid decarboxylase (EC 4.1.1.-).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Pred. No. 0.43;
                                modified and this statement is not removed.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               PROTEASE HOA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                31; Mismatches
                                                                                                                                                                   EMBL; AF114386; AAF24060.1; -...
EMBL; AL021710; CAA16717.1; ALT_SEQ.
EMBL; AL161548; CAB78839.1; ALT_SEQ.
MEROPS; S01.279; -...
InterPro; IPR001940; Protease2C,
InterPro; IPR001245; Ser_protease_Try.
Pfam; PF00089; trypsin; 1.
PRINTS; PR00834; PROTEASES2C.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Score 90;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           THYLAKOID.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                POLY-GLU.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                (Rel. 40, Created)
(Rel. 40, Last sequ
(Rel. 40, Last anno
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Pseudomonadaceae; Pseudomonas
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          34691 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               8.98;
                                                                                                                                                                                                                                                                                                                                                                                                                                                          Hydrolase; Serine protease;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            23.8%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    55; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Pseudomonas aeruginosa
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          321 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                NCBI_TaxID=287;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             16-OCT-2001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               PAAD_PSEAE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ACT_SITE
ACT_SITE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ACT_SITE
CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       TRANSIT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Local
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              TRANSIT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              PAAD_PSEAE
ρp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   THE STATE OF THE S
```

```
between the Saiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announce/or send an email to license@lsb-sib.ch).
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation -
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        92 NVDKDLVGWQAPQGSRSLTP----CTCGSSDL-----YLVTRHADVIPVRRGDS 137
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           91 QND-----WMAPPASGSSAPNAMVICPCSTGTLSAVATGACNNLIERAADVALKER---- 131
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           138 RGSLLSPR--PIS-----YLKGSSGGPLLCPAGHAVGIFRAAVSTRGVAKAVDFIPVES 189
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      80
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      29 QEEREVHFLISKAAQLVMATETDVALPAKPQAMQAFLTEYCGAAAGQI------RVFG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      43 QVEGEVQ-IVSTATQTFLATSINGVL------WIVYHGAGTRTIASPKGPVTQMYT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
                                                                                                                                                                                                                                                                                                                                                                                                                                         Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Kaneko T., Katoh T., Sato S., Nakamura A., Asamizu E., Tabata S.; "Structural analysis of Arabidopsis thaliana chromosome 3. II. Sequence features of the 4,251,695 bp regions covered by 90 Pl, TAC and BAC clones.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             STRAIN-cv. Columbia;
Kseelbach T., Bystedt M., Schroeder W.P.;
Submitted (JUL-2000) to the SWISS-PROT data bank.
-1- FUNCTION: SERINE PROTEASE THAT IS REQUIRED AT HIGH TEMPERATURE.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SEQUENCE FROM N.A., AND CHARACTERIZATION.
MEDLINE-98175982; PubMed-9507020;
Itzhaki H., Naveh L., Lindahl M., Cook M., Adam Z.;
"Identification and characterization of DegP, a serine protease associated with the luminal side of the thylakoid membrane.";
J. Biol. Chem. 273:7094-7098(1998).
                                                                                                                                                                                                                                                                                                                                                                                                                                      55;
                                                                                                                                                                                            EMBL; AE004818; AAG07406.1; -.
PIR, H83144; H83144.
InterPro; IPR003382; Flavoprotein.
Pfam; PF0441; Flavoprotein; I.
Hypothetical protein; Lyase; Decarboxylase; Complete proteome.
SEQUENCE 209 AA; 22367 MM; 01FD081CC495D3F6 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                    Length 209;
                                                                                                                                                                                                                                                                                                                                                                                                                                    61; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       DEGL_ARATH STANDARD; PRT; 437 AA.
022609; 09LK85;
16-CCT-2001 (Rel. 40, Created)
16-CCT-2001 (Rel. 40, Last sequence update)
28-FEB-2003 (Rel. 41, Last annotation update)
Protease Do-like 1, chloroplast precursor (EC 3.4.21.-).
DEGP1 OR DEGP OR ATSG27925 OR K16N12.18.
Arabidopsis thaliana (Mouse-ear cress).
                                                                                                                                                                                                                                                                                                                                                                                    DB 1;
                                                                                                                                                                                                                                                                                                                                                                                                         0.67;
                                                                                                                                                                                                                                                                                                                                                                                                                                 16; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                 Match 8.4%; Score 85.5; Local Similarity 27.9%; Pred. No. 0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SEQUENCE FROM N.A.
STRAIN-cv. Columbia;
MEDLINE-20363099; PubMed-10907853;
                                                                                                                                                                                                                                                                                                                                                                                                                                 51; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Res. 7:217-221(2000)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SEQUENCE OF 104-118
```

```
TISSUE-Brain
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            REPEAT
REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Matches
      SOURCE THE TENT TO DESCRIPTION OF THE TENT TO SOURCE OF THE TENT T
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         à
                                                                                                                                                                                            This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bloinformatics and the EMBL outstation the European Bloinformatics Institute. There are no restrictions on its most by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ..
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ------0A 102
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         103 PQGSRSLTPCTCGSSDLYLV-----TRHADVIPVRRRGDSRGSLLSPRPI 147
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                150 VPQGSGSGFVWDKQGHIVTNYHVIRGASDLRVTLADQTTFDAKVVGFDQDKDVAVLRIDA 209
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             210 PK -- NKLRPIPVGVSADLLVGQKVFAIGNPFGLDHTLTTGVISGLRREIS--SAATGRPI 265
                  CAN DEGRADE BETA-CASEIN.
ENZYME REGULATION: INHIBITED BY PHENYLMETHYLSULFONYL FLUORIDE AND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Transit peptide; Chloroplast; Thylakoid. CHLOROPLAST (POTENTIAL).
MAY BE INVOLVED IN THE DEGRADATION OF DAMAGED PROTEINS. IN VIVO,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    -> SSTLFLHSPPSSHL (IN REF
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SYL...--KGSSGGPLLCPAGHAVGIFRAAVSTRGVAKAVDF-IPVESL 190
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            266 QDVIQTDAAINPGNSGGPLLDSSGTLIGINTAIXSPSGASSGVGFSIPVDIV 317
                                                                             -i - SUBCELLULAR LOCATION: BOUND TO LUMINAL SIDE OF THE THYLAKOID
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      55;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CHARGE RELAY SYSTEM (POTENTIAL).
CHARGE RELAY SYSTEM (POTENTIAL).
CHARGE RELAY SYSTEM (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          8.2%; Score 83.5; DB 1; Length 437; 26.2%; Pred. No. 2.5;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          70 VYHGAGIRIIASPKGPVTQMY-------TNVDKDLVGW-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1497BlAB3F5FF2A4 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     2).

V -> I (IN REF. 2).

G -> R (IN REF. 2).

G -> D (IN REF. 2).

LL -> HF (IN REF. 2).

LL -> V (IN REF. 2).

L -> V (IN REF. 2).

L -> V (IN REF. 2).

G -> D (IN REF. 2).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    55;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            PROTEASE DO-LIKE 1. SERINE PROTEASE.
                                                                                                                     -1- INDUCTION: By heat shock.
-1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S2C
-1- SIMILARITY: Contains 1 PDZ/DHR domain.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Best Local Similarity 26.2%; Pred. No. 2.5; Matches 45; Conservative 17; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    HSPPSSQLSNST
                                                                                                                                                                                                                                                                                                                                                                                                                                  InterPro; 1901.1677.
InterPro; 1PR001478; PDZ.
InterPro; 1PR001254; Ser_protease_C.
InterPro; 1PR001254; Ser_protease_Try.
Pfam; PP00595; PDZ; 1.
Pfam; PP00899; trypsin; 1.
PRINTS; PR00834; PROTEASES_C.
SMART; SM00228; PDZ; 1.
PROSITE; PS50106; PDZ; 1.
Hydrolase; Serine protease; Transit pept
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            THYLAKOID
                                                                                                                                                                                                                                                                                                                                                       EMBL; AF028842; AAC39436.1; -.
EMBL; AP000371; BAB02539.1; -.
EMBL; AP001302; BAB02539.1; JOINED.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     36
38
38
38
38
38
38
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          103
437
321
421
171
201
280
                                                            O-PHENANTHROLINE
                                                                                                                                                                                                                                                                                                                                                                                                                        MEROPS; S01.279;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Hydrolase; STRANSIT TRANSIT CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ACT_SITE
ACT_SITE
ACT_SITE
CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         RESULT 12
8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss institute of Bioinformatics and the EMED cutstation the European Bioinformatics institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           68 WTVYHGAGTRIIASPKGPVTQMYINVDKDLVGWQAPQGSRSL----TPCTCGSSDLYLV 122
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         197 WMEWH-----PRAPVILLAGT-ADGNTWMWKVPNGDCKTFQGPNCPATCGR----- 240
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          123 TRHADVIPVRRR---GDSRGS-----1LSPRPISYLKGSSG--GPLLCPA----- 162
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           241 ----VLPDGKRAVVGYEDGTIRIWDLKQGSPIHVLKGTEGHQGPLTCVAANQDGSLILT 295
                                                                                                                                                                                                                                                              Gaps
                                                                                                                        Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostoml;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NCBL_TaxID-9606;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              -----GHAVGIFR----AAVSTRGVAKAVDFIPVESL 190
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          296 GSVDCQAKLVSATTGKVVGVFRPETVASQPSLGEGEESESNSVESL 341
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        HEPARIN-BINDING (POTENTIAL).
POLY-GLU.
WD 1.
WD 2.
WD 3.
WD 4.
WD 5.
WD 6.
WD 7.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            8.2%; Score 83; DB 1; Length 452;
25.3%; Pred. No. 2.9;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              DA1413D25EB236C0 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                47;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        EMBL; M95627; AAA68889.1; -.
PIR; I39383; I39383.
Genew, HGMC:18; AAMP.
MIN; 603488; -.
GO; GO:0008201; F:heparin binding activity; TAS.
InterPro; PR00180; WD40.
Fina; PF00400; WD40; 8.
SMART; SM03320; WD40; 8.
15-JUL-1998 (Rel. 36, Last sequence update)
28-FEB-2003 (Rel. 41, Last annotation update)
Angio-associated migratory cell protein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             13; Mismatches
                                                                                                                                                                                                                          SEQUENCE FROM N.A., AND PARTIAL SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           PROSITE; PSO0676; WD_REPEATS_1; 1.
PROSITE; PSSO082; WD_REPEATS_2; 6.
PROSITE; PSSO294; WD_REPEATS_REGION; 1
Repeat; WD_repeat.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              49015 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Conservative
                                                                                                    Homo sapiens (Human)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              452 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Local Similarity
ses 42; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          a
```

452 AA.

PRT;

AAMP\_HUMAN STANDARD; Q13685; 15-JUL-1998 (Rel. 36, Created)

AAMP\_HUMAN

```
VST2_HEVPA
P33426;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Signal.
SIGNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Matches
    SOUR PROPERTY NAME OF THE PROPERTY OF THE PROP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        d
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              113
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          165 CGGGGGSSSSSSCIHIWLVPGGTGNNGNCGCGGGGGSSSSSSCIHIKVENTDEQFL 224
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              54 AGSKLYATNGRLWEKELNGTGSWQKVSSSSVPTDSDK----KVWSIATDGNTFVLACVP 108
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              109 GTGVYKHCVNGAGSSSTGTTASPSTETCSQHAF----LVGGTSKPFWLVPGGTGNNGNCG 164
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ----- 151
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  63
                                                                                                                                                                                                                                                                                                                     Fraser C.M., Norris S.J., Weinstock G.M., White O., Sutton G.G., Dodson R., Gwinn M., Hickey E.K., Clayton R., Ketchum K.A., Sodergren E., Hardham J.M., McLeod M.P., Salzberg S., Peterson J., Khalak H., Richardson D., Howell J.K., Chidambaram M., Utterback T., McDonald L., Artiach P., Bowman C., Cotton M.D., Fujii C., Garland S., Wenter B., Horst K., Roberts K., Sandusky M., Weidman J., Smith H.O.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SGDTAYA-----QQTRGEQGTQKTSH----TGRDKNQVEGEVQIVSTATQTFLATSI-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              64 -NGVLWTVYHGAG---TRTIASPKGPVTQMYTNVDKDLVG-----WQAPQGSRSLTPCT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    71; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Science 281:375-388(1998).
-!- SUBCELLULAR LOCATION: Attached to the membrane by a lipid anchor
                                                                                                                                                                                       Spirochaetes; Spirochaetales; Spirochaetaceae; Treponema
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    HYPOTHETICAL LIPOPROTEIN TP0136. NACYL DIGLYCERIDE (POTENTIAL). GLY/SER RICH. GLY/SER-RICH. GLY/SER-RICH. POLY-SER. PICH. POLY-SER.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         *Complete genome sequence of Treponema pallidum, the syphilis
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        114 C----GSSDLYLVTRHADVIP------VRRRGDSRGSLLSPRPISYLK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (Potential).
-1. SIMILARITY: BELONGS TO THE TP013X FAMILY OF LIPOPROTEINS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    7.7%; Score 78.5; DB 1; Length 485; 23.4%; Pred. No. 8.3; tive 16; Mismatches 77; Indels 77
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        C7A4CEEDC7DC5CED CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          EMBL; AE001199; AAC65137.1; ALT_INIT.
IGR; TP0136; --
Hypothetical protein: Lipoprotein; Membrane; Signal;
Complete proteome.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     DMGEGYVVTTKHLYTKNGSSSAGPAQCPGGGGG 258
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      -----GSSGGPLLCPAGHAVG 167
                                           16-OCT-2001 (Rel. 40, Created)
16-OCT-2001 (Rel. 40, Last sequence update)
16-OCT-2001 (Rel. 40, Last annotation update)
Hypothetical lipoprotein TP0136 precursor.
485 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        POTENTIAL
                                                                                                                                                                                                                                                                                                       MEDLINE-98332770; PubMed-9665876;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           48984 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Conservative
  STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      23
485
24
178
210
267
327
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      485 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Query Match
Best Local Similarity
Matches 50; Conserv
                                                                                                                                                               Treponema pallidum
                                                                                                                                                                                                                                                      SEQUENCE FROM N.A.
                                                                                                                                                                                                           NCBI_TaxID-160;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    spirochete.
  Y136 TREPA
                                                                                                                                                                                       Bacteria;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  16
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      152
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    225
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               LIPID
DOMAIN
DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             DOMAIN
                                                                                                                                       TP0136
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   셤
ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation-the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   10;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     94 DKDLVGWQAPQGSRSLTPCTCGSSDLYLV -- - TRHADVIPVRRGDSRG - SLLSPRPISY 149
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                417 ENA----QQDKGIAIPHDIDLGESRVVIQDYDNQHEQDRPTPSPAPSRPFSVLRANDVLW 472
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                93
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   10 VGRINLSGDTAYAQQTRGEQGTQKTSHTGRDKNQV----EGEVQIVSTATQTFLA----T
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   -----RTIASPKG-PVTQMYTNV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      THAT IT MAY BE INVOLVED IN THE ENCAPSIDATION OF THE GENOMIC RNA BY EFFECTIVELY NEUTRALIZING THE NEGATIVELY CHARGED RNA.
                                                                                                                                                                                                                                                                                                                          MEDLINE-92024067; PubMed-1926770;
TIM A.W., Smith M.M., Guerra M.E., Huang C.-C., Bradley D.W.,
Fry K.E., Reyes G.R.;
"Hepatitis E virus (HEV): molecular cloning and sequencing of the (HI)-length viral genome.";
Virology 185:120-131(1991).
-i- FUNCTION: CONTAINS A HIGH BASIC AMINO ACID CONTENT SUGGESTING
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   57;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   150 LK------GSSGGPLLCPAGHAVGIFRAAVSTRGVAKAVDFIPV 187
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         7.7%; Score 78.5; DB 1; Length 660;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         473 LSLTAAEYDQSTYGSSTGPVY - VSDSVTLVNVATGAQAVARSLDWTKV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                85; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                POTENTIAL.
STRUCTURAL PROTEIN 2.
5832A013CCC4A61C CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             01-FBB-1994 (Rel. 28, Created)
01-FBB-1994 (Rel. 28, Last sequence update)
01-FBB-1994 (Rel. 28, Last annotation update)
Structural protein 2 precursor (ORF2).
Hepatitis E virus (strain Pakistan) (HEV).
Viruses: SSRNA positive-strand viruses, no DNA stage;
                                             01-DEC-1992 (Rel. 24, Created)
01-DEC-1992 (Rel. 24, Last sequence update)
01-DEC-1994 (Rel. 28, Last sequence update)
Structural protein 2 precursor (ORF2).
Hepatitis E virus (strain Burma) (HEV).
Viruses: ssRnA positive-strand viruses, no DNA stage;
Hepatitis E-like viruses.
Ş
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                42; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Pred. No. 12;
099
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                62 SINGV-----LWTVYHGAGT-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   660 AA; 70978 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       EMBL; M73218; AAA45736.1; -
PIR; C40778; VHWWH2.
InterPro; IPR004261; SP2.
Pfam; PF03014; SP2; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   19.78;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                45; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     STANDARD;
STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Local Similarity
                                                                                                                                                                                                                                                                                                SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        NCBI_TaxID=33774;
                                                                                                                                                                                                                                               NCBI_TaxID=31767;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   VST2_HEVPA
ID VST2_HI
DT 01-FEB
DT 01-FEB
DT 01-FEB
DE STRUCT
OS HEPATI
OC HEPATI
OC HEPATI
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its way non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          94 DKULVGWQAPQCSRSLTPCTCGSSDLYLV---TRHADVIPVRRRGDSRG-SILLSPRPISY 149
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   10 VGRINLSGDTAYAQQTRGEQGTQKTSHTGRDKNQV----EGEVQIVSTATQTFLA----T 61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Query Match
7.7%; Score 78.5; DB 1; Length 660;
Best Local Similarity 19.7%; Pred. No. 12;
Matches 45; Conservative 42; Mismatches 85; Indels 57; Gaps
                                                TSATEV S.A., EMETSON S.U., Reyes G.R., TSATEVA T.S., Legters L.J., Malik I.A., Iqbal M., Purcell R.H.;

"Characterization of a prototype strain of hepatitis E virus.";

Proc. Natl. Acad. Sci. U.S.A. 89:559-563(1992).

-I- FUNCTION: CONTAINS A HIGH BASIC AMINO ACID CONTENT SUGGESTING THAT IT MAY BE INVOLVED IN THE ENCAPSIDATION OF THE GENOMIC RNA BY EFFECTIVELY NEUTRALIZING THE NEGATIVELY CHARGED RNA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    150 LK------GSSGGPLLCPAGHAVGIFRAAVSTRGVAKAVDFIPV 187
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     LSLTAAEYDQSTYGSSTGPVY - - VSDSVTLVNVATGAOAVARSLDWTKV 519
                                                                                                                                                                                                                                                                                                                                                                                                                             1 22 BY SIMILARITY.
23 660 STRUCTURAL PROTEIN 2.
660 AA; 70980 MW; 8085BC53CFB46FD3 CRC64;
                                    MEDLINE-92115700; PubMed-1731327;
                                                                                                                                                                                                                                                                                                                                                   EMBL; M80581; AAA45727.1; -.
Interpro; IPR004261; SP2.
Pfam; PF03014; SP2; 1.
[1]
SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                  CHAIN
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        62
                                                                                                                                                                                                                                                                                                                                                                                                           Signal.
 q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 a
```

Search completed: August 30, 2003, 19:13:50 Job time: 10.7567 secs

hepatitis hepatitis

0991;75 0991;75 0991;75 0991;75 0991;75 0991;75 0991;75 0991;75 0991;75 0991;75 0991;75 0991;75 0991;75

091R83 091R08 091RR0 091RR0 091RR6 091RR5 09

hepatitis hepatitis hepatitis

hepatitis hepatitis hepatitis

hepatitis hepatitis hepatitis

nepatitis

hepatitis

hepatitis hepatitis

hepatitis

Run on:

```
(1)
SEQUENCE FROM N.A.
MEDLINE-2033484; PubMed=10864644;
Lai V.C., Zhong W., Skelton A., Ingravallo P., Vassilev V.,
Lai V.C., Zhong W., Skelton A., Ingravallo P., Vassilev V.,
Donis R.O., Rong Z., Lau J.Y.;
"Generation and characterization of a hepatitis C virus NS3 protease-dependent bovine viral diarrhea virus.";
J. Virol. 74:6339-6347(2000).
                                                                                                                                                                                                                                                                                                                                                                        Mucosal disease virus.
Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
Pestivírus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Lai V.C.H., Hong Z.;
Submitted (MAY-2000) to the EMBL/GenBank/DDBJ databases.
EMBL: A7268278; AAF82566.1; -.
HSSP: P26663; 1JXP.
MEROPS; S31.001; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        InterPro: JRR00128): CDVir_endptseP80.
InterPro: IRR001410; DRAD.
InterPro: IRR001410; DRAD.
InterPro: IRR001410; HV_NS3.
InterPro: IRR00156; HV_RARP.
InterPro: IRR001650; Helicase_C.
InterPro: IRR001659; RNA_pol_bs.
InterPro: IRR001695; RNA_pol_bs.
InterPro: IRR001095; RNA_pol_bs.
InterPro: IRR001094; RNA_pol_bs.
Fam: PF02907; HCV_NS3; 1.
Pfam: PF00291; Hclicase_C; 1.
Pfam: PF00998; Viral_RARP; 1.
                                                                                                                                                                                                                                                                                                                                     01-0CT-2000 (TrEMBLrel. 15, 01-0CT-2000 (TrEMBLrel. 15, 01-MAR-2003 (TrEMBLrel. 23, Genome polyprotein.
                                                                                                                                                                                                                                                                                                                     PRELIMINARY;
NCBI_TaxID-11099;
829
828
827.5
827.5
827
826.5
832
831
830.5
830
829.5
829
                                                                                                                                                                                                                                                                                                                    Q91FH8
Q91FH8;
                                                                                                                                                                                                                                                                                                  RESULT 1
Q91FH8
Q91fh8 mucosal dis
08579 hepatitis c
081756 hepatitis c
Q91fe5 hepatitis c
Q96188 hepatitis c
003463 hepatitis c
090wx5 hepatitis c
Q91wy hepatitis c
Q91rr8 hepatitis c
Q91rr5 hepatitis c
Q91rr5 hepatitis c
Q91rr7 hepatitis c
                                                                  (without alignments)
1352.314 Million cell updates/sec
                                                                                                                                                                                                                                                                                                                                                                                                                        Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.
                                                                                                             1 MKKKGSVVIVGRINLSGDTA.......VAKAVDFIPVESLETTMRSP 197
                                                          August 30, 2003, 19:00:22 ; Search time 37.5921 Seconds
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Description
      GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd
                                                                                                                                                                       Total number of hits satisfying chosen parameters:
                                                                                                                                                        830525 seqs, 258052604 residues
                                                                                                                                                                                                                                                                                                                                                                                                                                                            SUMMARIES
                                                                                                                                                                                                                  Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

    protein search, using sw model

                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Q9PWU9
Q91RR8
Q91RT5
Q91RR5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               036579
QBL756
Q9IFE5
Q9ELS8
Q03463
Q36608
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             091RR2
091RT9
091RR3
091RR4
                                                                                                                              BLOSUM62
Gapop 10.0 , Gapext 0.5
                                                                                                                                                                                                                                                            sp_archea:*
sp_bacteria:*
sp_fungl:*
sp_human:*
sp_invertebrate:*
sp_manmal:*
                                                                                                                                                                                                                                                                                                                                                                               sp_unclassified:*
                                                                                                                                                                                                                                                                                                                                                                      sp_vertebrate:*
                                                                                                                                                                                                                                                                                                                                                                                        sp_rvirus:*
sp_bacteriap:*
sp_archeap:*
                                                                                                                                                                                                                                                                                                                         sp_organelle:*
sp_phage:*
                                                                                                                                                                                         Minimum DB seq length: 0
Maximum DB seq length: 2000000000
                                                                                                                                                                                                                                                                                                                                                    sp_rodent:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       US-09-965-594-22
1016
                                                                                                                                                                                                                                                                                                                                           sp_plant:*
                                                                                                                                                                                                                                                                                                                                                              sp_virus:*
                                                                                                                                                                                                                                                     SPTREMBL_23:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       2436
3011
3011
3011
3011
3015
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   181
181
181
181
181
181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Query
Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     862.5
844.5
844.5
844.5
844.5
843.5
841.5
841.5
                                                                                            Title:
Perfect score:
                                                                                                                               Scoring table:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Score
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           837
837
837
836
836
                                         OM protein
                                                                                                              Sequence:
                                                                                                                                                         Searched:
                                                                                                                                                                                                                                                     Database
```

654821098765482P

Result

Last sequence update) Last annotation update)

Created)

PRT; 4040 AA

ALIGNMENTS

```
Query Match
Pfam;
                       Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 RESULT 3
Q81756
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     d
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ద
                                                                                                                                                                                                                                                                      ï
                                                                                                                                                                                                                                                                                                                                                                                                                62 SINGVLWTVYHGAGTRTIASPKGPVTQMYTNVDKDLVGWQAPQGSRSLTPCTCGSSDLYL, 121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                           122 VTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVSTRGVAKA 181
                                                                                                                                                                                                                                                                                                                                       10 GSVVIVGRIVLSGSGSITACAQQTRGLIGCKITSLTGRDKNQVEGEVQIVSTATQTFLAT 69
                                                                                                                                                                                                                                                                                                                 61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             LTABLE THE VICTOR IN THE VICTOR OF THIS VICTOR OF THE VICTOR OF THE CONTRINGUENCE OF THE SCIENCE 277:570-574(1997).

-!- SUBUNIT: THE VIRTON OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MRNA (BY SIMILARITY).

EMBL: AR009606; AAB66324.1; --
HSSP: P27958; 1HEI.
                                                                                                                                                                                                                                                                                                               5 GSVVIVGRINLSGD---TAYAQQTRGEQGTQKTSHTGRDKNQVEGEVQIVSTATQTFLAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                S.M.
                                                                                                   PROSITE; P550521; RDRP_VIRAL; 1.
PROSITE; P500531; RNASE_T2_2; 1.
ATP-binding; Helicase; Hydrolase; Nonstructural protein; Polyprotein; RNA-directed RNA polymerase; Transferase.
SEQUENCE 4040 AA; 453073 MW: ADE87791D055B9DC CRC64;
                                                                                                                                                                                                                                                                        3; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
Hepacivirus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Kolykhalov A.A., Agapov E.V., Blight K.J., Mihalik K., Feinstone Rice C.M.;
                                                                                                                                                                                                                          Ouery Match 84.9%; Score 862.5; DB 12; Length 4040; Best Local Similarity 88.7%; Pred. No. 1.9e-72; Matches 173; Conservative 5; Mismatches 14; Indels 3;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    "Transmission of hepatitis C by intrahepatic inoculation with
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           01-JAN-1998 (TrEMBLrel. 05, Created)
01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          PRT; 3011 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Helicase_C.
RNA_pol_DS_PS.
RNA_pol_PSvir.
                 SMART; SM00487; DEXDC; 1.
SMART; SM00490; HELLCC; 1.
PROSITE; PS00037; MYB_1; 1.
PROSITE; PS50507; RDRP_POSITIVE; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Interpro; IPRU02223, HCV_NS...
Interpro; IPR002531; HCV_NS2.
Interpro; IPR002518; HCV_NS2.
***erpro; IPR004109; HCV_NS3.
***erpro; IPR004109; HCV_NS3.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            HCV_capsid.
HCV_core.
HCV_env.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          MEDLINE-97373636; PubMed-9228008
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           HCV_NS4b.
HCV_NS5a.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  HCV_RdRP
  PRINTS; PR00729; CDVENDOPTASE.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           182 VDFIPVESLETIMRS 196
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   InterPro; IPR001410; DEAD.
InterPro; IPR002522; HCV_C
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         IPR007095; I
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 IPR002521;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     IPR000745;
IPR001490;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          IPR001650;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  IPR002166;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 IPR002868
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Genome polyprotein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SEQUENCE FROM N.A. STRAIN-H77;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          NCBI_TaxID=11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              InterPro;
InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             InterPro;
InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        036579
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              036579
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          DER PER PRESENTATION OF THE PRESENTATION OF TH
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  q
    Q
                                                                                                                                                                                                                                                                                                                                                                                                                             ď
                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  q
                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
```

```
1064
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               113
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  173
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   23
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               114 CGSSDLYLVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   3 KKGSVVIVGRIN-----LSGDTAYAQQTRGEQGTQKTSHTGRDKNQVEGEVQIVST
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            9; Gaps
                                                                                                                                                                                                                                                                                                                                                                      ATP-binding; Coat protein; Envelope protein; Glycoprotein; Helicase;
Hydrolase; Nonstructural protein; Polyprotein;
RNA-directed RNA polymerase; Transferase; Transmembrane.
SEQOENCE 3011 AA; 327182 MM; E2E0EE809GG3C189 CRG64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         DB 12; Length 3011;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            SEQUENCE FROM N.A.
Choo Q.-L., Richman K., Han J.;
Choo Q.-L., Richman K., Han J.;
The nucleotide sequence of the Hepatitis C viral genome.*;
Submitted (MAY-1990) to the EMBL/GenBank/DDBJ databases.
EMBL; M32084; AAA45677.1; -.
HSSP; P27958; 1A1V.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    83.5%; Score 848.5; DB 12; Length
82.4%; Pred. No. 2.8e-71;
.ive 10; Mismatches 17; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Q81756 PRELIMINARY; PRT; 2436 AA. 081755; CITEMBLED. 01, Created) 01-NOV-1996 (TrEMBLED. 01, Last sequence update) 01-NOR-2003 (TrEMBLED. 23, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        InterPro; IPR001410; DEAD.
InterPro; IPR002531; HCV_NS1.
InterPro; IPR002518; HCV_NS2.
InterPro; IPR004109; HCV_NS3.
InterPro; IPR001409; HCV_NS4.
InterPro; IPR001409; HCV_NS4b.
InterPro; IPR001409; HCV_NS4b.
InterPro; IPR001569; HCV_RGAB.
InterPro; IPR001569; HCV_RGAB.
InterPro; IPR001569; HCV_RGAB.
InterPro; IPR001059; Halicase_C.
InterPro; IPR001094; NNA_POL_DS_PS.
InterPro; IPR001094; NNA_POL_PSyir.
Pfam; PP01560; HCV_NS1; I.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            174 STRGVAKAVDFIPVESLETTMRSP 197
                                                                                                                                                                                                                                                                                       SMART; SM00487; DEXDC; 1.
PROSITE; PS50507; RDRP_POSITIVE; 1.
PROSITE; PS50521; RDRP_VIRAL; 1.
R Pfam, PF01539; HCV_env; 1.
R Pfam, PF01560; HCV_env; 1.
R Pfam, PF01581; HCV_NS2; 1.
R Pfam, PF02907; HCV_NS3; 1.
Pfam, PF01001; HCV_NS4; 1.
Pfam, PF01001; HCV_NS4b; 1.
Pfam, PF01001; HCV_NS4b; 1.
Pfam, PF01001; HCV_NS4b; 1.
Pfam, PF00271; helicase_C; 1.
Pfam; PF00271; helicase_C; 1.
Pfam; PF00998; Viral_KGRP; 1.
ProDom; PD186062; HCV_NS1; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Genome polyprotein (Fragment). Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      82.4%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Best Local Similarity 82.4%
Matches 168; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Pfam; PF01560; HCV_NS1; 1
Pfam; PF01538; HCV_NS2; 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Hepacivirus.
NCBI_TaxID=11103;
```

Ĥ

```
09ELS8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          RESULT 5
09ELS8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          DE REPRESENTATION OF COLUMN AND DESTRUCTURA DE REPRESENTATION DE R
      SKKKK PREPARATION OF THE STATE 
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                614
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             54 ATQTFLATSINGVLWTVYHGAGTRTIASPKGPVTQMYTNVDKDLVGWQAPQGSRSLTPCT 113
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       674
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    173
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 53
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CGSSDLYLVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      3 KKGSVVIVGRIN-----LSGDTAYAQQTRGEQGTQKTSHTGRDKNQVEGEVQIVST
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           J. Gen. Virol. 82:1291-1297(2001).

-!- SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLECCAPSID COVERED BY A LIPOPMOTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
PROTEIN H AND GLYCOPROTEIN E. THE NUCLECCAPSID IS A COMPLEX OF PROTEIN C AND MRNA (BY SIMILARITY).

EMBL. AF21543; AAF81759.1; -...
INTERPRO; IPR000345; CytC_heme_bind.

InterPro; IPR00140; DEAD.
                                                                                                                                                                       SMART; SMO487; DEXDC; 1.
PROSITE; PS50507; RDRP_POSITIVE; 1.
PROSITE; PS50507; RDRP_POSITIVE; 1.
APP-binding; Coat protein; Envelope protein; Glycoprotein; Helicaso; Hydrolase; Nonstructural protein; Polyprotein;
RNA-directed RNA Polymerase; Transferase; Transmembrane.
NON_IER
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ENGUENCE FROM N.A.
MEDLINE-21262212; PubMed-11369872;
Lanford R.E., Lee H., Chavez D., Guerra B., Brasky K.M.;
"Infectious cDNA clone of the hepatitis C virus genotype 1 prototype
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Viruses: ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
Hepacivirus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        83.1%; Score 844.5; DB 12; Length 2436; 82.4%; Pred. No. 5.1e-71; tive 9; Mismatches 18; Indels 9;
                                                                                                                                                                                                                                                                                                                                                                                                                                                           2436 AA; 264734 MW; D7B9872900BE3125 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            01-0cT-2000 (TrEMBLrel. 15, Created)
01-0cT-2000 (TrEMBLrel. 15, Last sequence update)
01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      PRT; 3011 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        174 STRGVAKAVDFIPVESLETTMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  735 CTRGVAKAVDFIPVENLETTMRSP 758
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      InterPro; IPR002522; HCV_capsid.
InterPro; IPR002521; HCV_core.
InterPro; IPR002519; HCV_env.
Pfam; PF01006; HCV_NS4a; 1.
Pfam; PF01001; HCV_NS4a; 1.
Pfam; PF01001; HCV_NS5a; 1.
Pfam; PF0021; HCV_NS5a; 1.
Pfam; PF00299; Viral_RGRP; 1.
ProDom; PD186062; HCV_NS1; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       HCV_NS3.
HCV_NS4a.
HCV_NS4b.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          InterPro; IPR002531; HCV_NS1.
InterPro; IPR002518; HCV_NS2.
InterPro; IPR004109; HCV_NS3.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Best Local Similarity 82.4
Matches 168; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                  2436
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          InterPro; IPR000745;
InterPro; IPR001490;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Genome polyprotein.
Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    NCBI_TaxID-11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                  2436
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           sednence.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    114
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      O9IFES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        õ
```

```
54 ATQTFLATSINGVLWTVYHGAGTRTIASPKGPVTQMYTNVDKDLVGWQAPQGSRSLTPCT 113
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            3 KKGSVVIVGRIN-----LSGDTAYAQQTRGEQGTQKTSHTGRDKNQVEGEVQIVST
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Desai S.W., Devare S., Yamaguchi J.;
"Hepatitis C Virus.";
Submitted (JUL-2000) to the EMBL/GenBank/DDBJ databases.
-I-SUBUNIT: THE VIRION OF THIS VIROS IS A NUCLECCAPSID COVERED BY A LIPOPROTEIN ENVELOPE CONSIGNS OF TWO PROTEINS:
PROTEIN M AND GRYCOPROTEIN E. THE NUCLECCAPSID IS A COMPLEX OF PROTEIN C AND MARA (PY SIMILARITY).
EMBL: AF200978; AAG2099.1; --
HSSP: P27985; JHEI.
InterPro; IPR000345; CytC_heme_bind.
InterPro; IPR001410; DEAD.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                            SMART; SM00487; DEXDC. 1.
PROSITE; PS00190; CYTOCHRONE_C; 1.
PROSITE; PS50507; RDRP_POSITIVE; 1.
ATP-Dinding; Coat protein; Envelope protein; Glycoprotein; Helicase; Hydrolase; Nonstructural protein; Polyprotein;
RNA-directed RNA polymerase; Transmembrane.
SEQUENCE 3011 AA; 327124 MW; 2489CE74AC864E58 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 9; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           01-MAR-2001 (TrEMBLrel. 16, Created)
01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
Genome polyprotein.
18-partial C virus.
Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae; Hepatius.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             DB 12; Length 3011;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Query Match 83.1%; Score 844.5; DB 12; Length Best Local Similarity 82.4%; Pred. No. 6.7e-71; Matches 168; Conservative 9; Mismatches 18; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         PRT; 3011 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1185 CTRGVAKAVDFIPVENLETTMRSP 1208
InterPro; IPR001166; HCV_RdRP.
InterPro; IPR001165; Helicase_C.
InterPro; IPR001050; Helicase_C.
InterPro; IPR001094; RNA_pol_DS_PS.
InterPro; IPR007094; HCV_core; 1.
Pfam; PP01559; HCV_CORE; 1.
Pfam; PP01006; HCV_NS2; 1.
Pfam; PP01006; HCV_NS3; 1.
Pfam; PP01001; HCV_NS4; 1.
Pfam; PP01001; HCV_NS4; 1.
Pfam; PP01001; HCV_NS4; 1.
Pfam; PP00998; Viral_RdRP; 1.
Pfam; PP00998; Viral_RdRP; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                174 STRGVAKAVDFIPVESLETTMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             NCBI_TaxID=11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    STRAIN-colonel;
```

InterPro; InterPro; InterPro; InterPro; InterPro;

nterPro;

InterPro; InterPro; InterPro; interPro;

InterPro;

```
Okamoto H., Okada S., Sugiyama Y., Kurai K., Iizuka H., Machida A., Miyakawa Y., Mayumi M.;
"Nucleotide sequences of the genomic RNA of hepatitis C virus isolated from a human carrier: comparison with reported isolates for conserved and divergent regions.";
J. Gen. Virol. 72:2697-2704(1991).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Okamoto H., Kanai N., Mishiro S.;
"Full-length nucleotide sequence of a Japanese hepatitis C virus isolate (HC-JI) with high homology to USA isolates.";
Nucleic Acids Res. 20:6410-6410(1992).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            PROSITE: PSS0507; RDRP POSITIVE: 1.
PROSITE; PSS0521; RDRP_VIRAL; 1.
ATP-binding; Coat protein; Envelope protein; Glycoprotein;
Hydrolase; Monstructural protein; Polyprotein;
"The 5'-terminal sequence of the hepatitis C virus genome. Jpn. J. Exp. Med. 60:167-177(1990). [2] SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      MEDLINE-94174722; PubMed-7510436;
Mink M., Benichou S., Madaule P., Tiollais P., Prince A.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Hydrolase; Nonstructural protein; rosyrocom; RNA-directed RNA polymerase; Transferase; Transmembrane-cennewce 3011 AA; 327112 MW; 97E9052C0250463B CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Submitted (DEC-1992) to the EMBL/GenBank/DDBJ databases
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          InterPro: IPR002519; HCV_env.
InterPro: IPR002519; HCV_env.
InterPro: IPR002519; HCV_NS1.
InterPro: IPR004109; HCV_NS2.
InterPro: IPR004109; HCV_NS4.
InterPro: IPR000445; HCV_NS4b.
InterPro: IPR001490; HCV_NS4b.
InterPro: IPR00166; HCV_RGAB.
InterPro: IPR00166; HCV_RGRP.
InterPro: IPR00166; HCV_RGRP.
InterPro: IPR001650; Helicase_C.
InterPro: IPR007095; RNA_POL_DS_PS.
InterPro: IPR007094; RNA_POL_DS_PS.
IPR0070950; HCV_COTE; I.
                                                                                                                                      MEDLINE-9204440; PubMed-1658196;
                                                                                                                                                                                                                                                                                                                                                                                                       MEDLINE-93117120; PubMed-1335573;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   HCV_core.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Pfam; PF01006; HCV_NS4a; 1.
Pfam; PF01001; HCV_NS4b; 1.
Pfam; PF01506; HCV_NS5a; 1.
Pfam; PF000271; holicase_C; 1.
Pfam; PF00998; Viral_RGRP; 1.
ProDom; PD186062; HCV_NS1; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    InterPro; IPR001410; DEAD.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                HCV_NS1; 1
HCV_NS2; 1
HCV_NS3; 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SM00487; DEXDC;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   InterPro; IPR002522;
InterPro; IPR002521;
                                                                                                                                                                                                                                                                                                                                                       SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Pfam; PF02907; H
Pfam; PF01006; H
Pfam; PF01001; H
Pfam; PF01506; H
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    PF01539; PF01560; P
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          PF01538;
PF02907;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  STRAIN-HC-J1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Inchauspe G.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   STRAIN-HC-J]
                                                                                                                                                                                                                                                                                                                                                                                STRAIN-HC-J]
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Okamoto H
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SMART;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Pfam;
Pfam;
     SOW WENT THE TRANSPORT OF THE TRANSPORT 
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ATQTFLATSINGVLWTVYHGAGTRTIASPKGPVTQMYTNVDKDLVGWQAPQGSRSLTPCT 113
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      114 CGSSDLYLVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAV 173
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          23
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       3 KKGSVVIVGRIN-----LSGDTAYAQQTRGEQGTQKTSHTGRDKNQVEGEVQIVST
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     9; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     protein; Glycoprotein; Helicase;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   DB 12; Length 3011;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Okamoto H., Okada S., Sugiyama Y., Yotsumoto S., Tanaka T.,
Yoshizawa H.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             83.1%; Score 844.5; DB 12; Length
81.9%; Pred. No. 6.7e-71;
.ive 11; Mismatches 17; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 PROSITE; PS00190; CYTCCHROME_C; 1.
PROSITE; PS00190; CYTCCHROME_C; 1.
PROSITE: PS50507; RORP_POSITIVE; 1.
ATP-binding; Coat protein; Envelope protein; Glycoprotein Hydrolase; Monstructural protein; Polyprotein; RNA-directed RNA-polymerase; Transferase; Transmembrane, SEQUENCE 3011 AA; 327107 MW; AGBECFSA1B3EE11F CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              003463 PRELIMINARY; PRT; 3011 AA. 003463; CITEMELEI. 01, Created) 01-NOY-1996 (TIEMELEI. 01, Last sequence update) 01-NOY-1996 (TIEMELEI. 23, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       STRGVAKAVDFIPVESLETIMRSP 197
                                                                                                                                                                                                                 IPR002868; HCV_NSSA.
IPR001166; HCV_RGRP.
IPR001169; Helicase_C.
IPR007095; RNA_pol_DS_PS.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     MEDLINE-91013116; PubMed-2170712;
                                                                                                          IPR002518; HCV_NS2.
IPR004109; HCV_NS3.
IPR000745; HCV_NS4a.
IPR001490; HCV_NS4b.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Pfam; PF02907; HCV_NS3; 1.
Pfam; PF02090; HCV_NS4s; 1.
Pfam; PF01001; HCV_NS4s; 1.
Pfam; PF01506; HCV_NS5s; 1.
Pfam; PF001506; HCV_NS5s; 1.
Pfam; PF001998; Viral_RGRP; 1.
Probom; PD186062; HCV_NS1; 1.
                                                                                                                                                                                                                                                                                                                                                  Pfam; PF01543; HCV_capsid: 1
Pfam; PF01542; HCV_core; 1.
Pfam; PF01559; HCV_env; 1.
Pfam; PF01560; HCV_NS1; 1.
Pfam; PF01558; HCV_NS2; 1.
Pfam; PF02907; HCV_NS3; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Best Local Similarity 81.9
Matches 167; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                SM00487; DEXDC;
                                                                                      PR002531;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Genome polyprotein.
Hepatitis C virus.
                                 PR002521
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Hepaclvirus.
NCBI_TaxID=11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            STRAIN-HC-J]
```

Query Match

ŏ qq ò g à g ŏ q

174

RESULT (
003463
AC 0003463
AC 0000
DT 011 001
DT 011 001
DT 011 001
DT 011 011
DT 011 01

Helicase;

S

```
IPR002166;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Hepacivirus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (nterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 interPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       O9PWX5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Q9PWX5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           œ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           RESULT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   09PWX5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     DER PRESENTATION OF THE PROPERTY OF THE PROPER
DR KEE
                                                                                                                                                                                                                                                                                                                                qq
                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          요
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                          1005 RKGREILLGPADGMYSKGWRLLAPITAYAQQTRGLLGCIITSLTGRDKNQVEGEVQIVST 1064
                                                                                                                                                                                      54 ATQTFLATSINGVLWTVYHGAGTRTIASPKGPVTQMYTNVDKDLVGWQAPQGSRSLTPCT 113
                                                                                                                                                                                                                                                                         114 CGSSDLYLVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAV 173
                                                                                                      53
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     MEDLINE-97385173; PubMed-9238047; Yanaqi M., Purcell R., Emerson S.U., Bukh J.; Yanaqi M., Purcell R., Emerson S.U., Bukh J.; "Transcripts from a single full-length cDNA clone of hepatitis C virus are infectious when directly transfected into the liver of a
                                                                                                      3 KKGSVVIVGRIN-----LSGDTAYAQQTRGEQGTQKTSHTGRDKNQVEGEVQIVST
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Genome polyprotein.
Hepatitis C virus strain H77.
Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Proc. Natl. Acad. Sci. U.S.A. 94:8738-8743(1997).
Proc. Natl. Acad. Sci. U.S.A. 94:8738-8743(1997).
LYOPEROTE THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY LIPOPROTEIN ENVELOPE. CHE ENVELOPE CONSISTS OF TWO PROTEINS:
PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C. AND MRNA (BY SIMILARITY).
                    DB 12; Length 3011;
                                                            6
                                                            Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  01-JAN-1998 (TrEMBLrel. 05, Created)
01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
                                                         19;
               Score 843.5; DB 17
Pred. No. 8.4e-71;
8; Mismatches 19.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            PRT; 3011 A.A.
                                                                                                                                                                                                                                                                                                                                                                             Interpro; IPR001410; DEAD.
Interpro; IPR00122; HCV_capsid.
Interpro; IPR00252; HCV_capsid.
Interpro; IPR002519; HCV_core.
Interpro; IPR002519; HCV_env.
Interpro; IPR002519; HCV_env.
Interpro; IPR002519; HCV_NS2.
Interpro; IPR001409; HCV_NS3.
Interpro; IPR001409; HCV_NS4b.
Interpro; IPR001409; HCV_NS4b.
Interpro; IPR001409; HCV_NS4b.
Interpro; IPR001409; HCV_NS4b.
Interpro; IPR001569; HCV_NS4b.
Interpro; IPR001569; HCV_NS4b.
Interpro; IPR001569; HCV_NS5a.
Interpro; IPR001569; HCV_NS4b.
Interpro; IPR001704; RNA_POI_DS_PSVIT.
Pfam: PF01543; HCV_Capsid; 1.
                                                                                                                                                                                                                                                                                                                                                         174 STRGVAKAVDFIPVESLETIMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                EMBL; AF011751; AAB67036.1; -. HSSP; P27958; 1HEI.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            HCV_core; 1.
HCV_env; 1.
HCV_NS1; 1.
HCV_NS2; 1.
HCV_NS3; 1.
HCV_NS4a; 1.
               83.08;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 HCV_NS5a; 1.
                                        82.48;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               HCV_NS4b;
                                        Best_Local Similarity 82.4
Matches 168; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Hepacivirus.
NCBI_TaxID=63746;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   PF01506; PF00271; )
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                PF01542;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         PF01539;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          PF01560;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   PF01538;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         PF01006;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      STRAIN-H77
                      Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          036608
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Pfam;
Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Pfam;
Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                    RESULT OSE608
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                       g
                                                                                                                                                                                           ò
                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                         õ
                                                                                                                                                                                                                                                                                                                                                           ģ
```

```
ij
                                                                                                                                                                                                                                                                                                                         1005 RRGQEILLGPADGMVSKGWRLLAPITAYAQQIRGLLGCIITSLIGRDKNQVEGEVQIVST 1064
                                                                                                                                                                                                                                                                                         53
                                                                                                                                                                                                                                                                                                                                                                       54 ATQTFLATSINGVLWTVYHGAGTRTIASPKGPVTQMYTNVDKDLVGWQAPQGSRSLTPCT
                                                                                                                                                                                                                                                                                                                                                                                                                                                       114 CGSSDLYLVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAV
                                                                                                                                                                                                                                                                                       3 KKGSVVIVGRIN-----LSGDTAYAQQTRGEQGTQKTSHIGRDKNQVEGEVQIVST
                                                                                                                                                                                                                                             9; Gaps
                                                      PROSITE; PSSOSOT: RDRP_POSITIVE; 1.
PROSITE; PSSOSOT: RDRP_VIRAL: 1.
ATP-blading; Coat protein; Browelope protein; Glycoprotein; Helicase;
Hydrolase; Nonstructural protein; Polyprotein;
RNA-directed RNA polymerase; Transferase; Transmembrane.
SEQUENCE 3011 AA: 327112 MW; 0B75E6B81CB5C198 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         MEDLINE-99420396; PubMed-10489358;
Yanaqi M., Purcell R.H., Emerson S.U., Bukh J.;
"Hepatitis C virus: an infectious molecular clone of a second major genotype (2a) and lack of viability of intertypic la and 2a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Viruses; saRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                         DB 12; Length 3011;
                                                                                                                                                                                                  Query Match 82.8%; Score 841.5; DB 12; Length Best Local Similarity 81.9%; Pred. No. 1.3e-70; Matches 167; Conservative 10; Mismatches 18; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Last sequence update)
Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     PRT; 3015 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              174 STRGVAKAVDFIPVESLETTMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          01-MAY-2000 (TIEMBLrel. 13, Created)
01-MAY-2000 (TIEMBLrel. 13, Last sequ
01-MAR-2003 (TIEMBLrel. 23, Last anno
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        HCV_capsid.
HCV_core.
HCV_env.
HCV_NSI.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          HCV_NS2.
HCV_NS3.
HCV_NS4a.
HCV_NS4b.
HCV_NS5b.
HCV_NS5a.
Pfam; PF00998; Viral_RdRP; 1.
ProDom; PD186062; HCV_NS1; 1.
SMART; SM00487; DEXDC; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Virology 262:250-263(1999).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   InterPro; IPR001410; DEAD.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              IPR004109;
IPR000745;
IPR001490;
IPR002868;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              InterPro; IPR002521;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       IPR002531;
IPR002518;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      IPR002519;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Genome polyprotein.
Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     NCB1_TaxID-11103;
```

```
Q91RR8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               091RR8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Best Loc
Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           RESULT 10
Q91RR8
        õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      엄
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ä
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1128
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        54 ATQTFLATSINGVLWTVYHGAGTRTIASPKGPVTQMYTNVDKDLVGWQAPQGSRSI,TPCT 113
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  53
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  3 KKGSVVIVGRIN-----LSGDTAYAQQTRGEQGTQKTSHTGRDKNQVEGEVQIVST
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Submitted (AUG-1999) to the EMBL/GenBank/DDBJ databases.
-!- SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS: PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        9; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                   PROSITE; PS00392; DDC_GAD_HDC_YDC; 1.
PROSITE; PS50507; RDRP_POSITIVE; 1.
ATP-binding; Coat protein; Envelope protein; Glycoprotein; Helicase;
Hydrolase; Nonstructural protein; Polyprotein;
RNA-directed RNA polymerase; Transferase; Transmembrane.
SEQUENCE 3015 AA; 328159 MW; B7D23BC1F190665A CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                01-MAY-2000 (TrEMBLrel. 13, Created)
01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
01-MAY-2003 (TrEMBLrel. 23, Last annotation update)
Genome polyprotein.
Hepatitis C virus.
Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                DB 12; Length 3015;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                82.8%; Score 841.5; DB 12; Length
81.9%; Pred. No. 1.3e-70;
Live 10; Mismatches 18; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          PRT; 3015 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          174 STRGVAKAVDFIPVESLETTMRSP 197
                             InterPro; IPR002129; Pyridoxal_deC.
InterPro; IPR007095; RNA_pol_DS_PS.
InterPro; IPR007094; RNA_pol_PSvir
  InterPro; IPR001650; Helicase_C
                                                                                                        Pfam; PF01543; HCV_caps.d. 1. Pfam; PF01542; HCV_care; 1. Pfam; PF01539; HCV_env; 1. Pfam; PF01539; HCV_env; 1. Pfam; PF01539; HCV_NS; 1. Pfam; PF01560; HCV_NS; 1. Pfam; PF01001; HCV_NS49; 1. Pfam; PF01001; HCV_NS49; 1. Pfam; PF01506; HCV_NS49; 1. Pfam; PF01506; HCV_NS59; 1. Pfam; PF01506; HCV_NS59; 1. Pfam; PF00221; helicase_C: 1. Pf
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Best Local Similarity 81.9
Matches 167; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                  DEXDc; ]
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         NCBI_TaxID=11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                  SM00487;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Hepacivirus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       09PWU9;
                                                                                                                                                                                                                                                                                                                                                                                                                                                    SMART;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       QUARTU9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   RESULT 9
DOR NO DOR NEW KEE NO DOR NO DOR NO DOR NEW KEE NO DOR N
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ΩD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
```

```
1129 GGSSDLYLVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGLFRAAV 1188
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         114 CGSSDLYLVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAV 173
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1009 RRGQEILLGPADGMVSKGWRLLAPITAYAQQTRGLLGCIITSLTGRDKNQVEGEVQIVST
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        3 KKGSVVIVGRIN-----LSGDTAYAQQTRGEQGTQKTSHTGRDKNQVEGEVQIVST
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                01-DEC-2001 (TrEMBLrel. 19, Created)
01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
NS3 protease (Fragment).
Hepatitis C virus.
Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               82.8%; Score 841.5; DB 12; Length 3015;
81.9%; Pred. No. 1.3e-70;
ive 10; Mismatches 18; Indels 9;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        PROSITE; PS00392; DEC.GAD_HDC_VDC; 1.
PROSITE; PS50507; RDRP_POSITIVE; 1.
PROSITE; PS50521; RDRP_VRAL; 1.
ATP-binding; Cast protein; Envelope protein; Glycoprotein; Ilydrolase; Nonstructural protein; Polyprotein;
Rydrolase; Nonstructural protein; Polyprotein;
Rydricted RNA polymerase; Transferase; Transmembrane.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           328084 MW; E309F6318067D6CD CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  181 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1189 CTRGVARAVDFIPVENLGTTMRSP 1212
                                                                                                                                                                                                                                                                                        InterPro: IPR001499; HCV_NS4D:
InterPro: IPR001668; HCV_NS5A:
InterPro: IPR001666; HCV_NS5A:
InterPro: IPR001650; HCLTGASC.
InterPro: IPR001059; Pyridoxal_dec.
InterPro: IPR007059; RNA_POL_DS_PS.
InterPro: IPR007094; RNA_POL_DS_PS.
InterPro: IPR007094; RNA_POL_DS_PS.
InterPro: IPR007094; RNA_POL_PS_PS.
InterPro: IPR007094; RNA_POL_PS_PS.
InterPro: IPR007094; RNA_POL_PS_PS.
IPfam: PP01560; HCV_Core; I.
Pfam: PP01539; HCV_Core; I.
Pfam: PP01006; HCV_NS1; I.
Pfam: PP01006; HCV_NS1; I.
Pfam: PP01001; HCV_NS3; I.
Pfam: PP01001; HCV_NS4b; I.
Pfam: PP01001; HCV_NS4b; I.
Pfam: PP01001; HCV_NS4b; I.
Pfam: PP00271; helicase C. I.
Pfam: PP00271; helicase C. I.
Propon: PD186662; HCV_NS1; I.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      174 STRGVAKAVDFIPVESLETTMRSP 197
EMBL; AF177039; AAF01181.1; -. EMBL; AF177037; AAF01179.1; -.
                                                                                IPR001410; DEAD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Local Similarity 81.9
hes 167; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       DEXDc;
                                                                                                         IPR002522; H
IPR002521; H
IPR002519; H
                                                                                                                                                                                         IPR002531;
IPR002518;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SEQUENCE 3015 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Hepacivirus.
NCBI_TaxID=11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SMART; SM00487;
                                                                                                                                   InterPro;
InterPro;
                                                                                                                                                                                                                InterPro;
InterPro;
                                                                             interPro;
                                                                                                            InterPro;
                                                                                                                                                                                         InterPro;
                                                                                                                                                                                                                                                                         InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Query Match
```

113

ä

~

```
79 IASPKGPVTQMYTNVDKDLVGWQAPQGSRSLIPCTCGSSDLYLVTRHADVIPVRRKGDSR 138
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   4 TAYAQOTRGLLGCIXTSLTGRDKNQVEGEVQIVSTAAQIFLATCINGVCWTVYHGAGTRT 63
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             19 TAYAQQIRGEQGTQKTSHTGRDKNQVEGEVQIVSTAIQIFLAISINGVLWTVYHGAGIRI 78
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 139 GSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVSTRGVAKAVDFIPVESLETTMRS 196
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        124 GSLLSPRPISYLKGSSGGPLLCPAGHAVGIFKAAVCTRGVAKAVDFIPVENLETTMRS 181
                                                                                                                                                                                                                                                                                                                                                                                 Holland-Staley C.A., Kovari L.C., Golenberg E., Mayers D.L.;
"Genetic Diversity and response to IFN of the NS3 Protease Gene from
Clinical Strains of the Hepatitis C Virus.";
EMBL, Genetic (APR-2001) to the EMBL/GenBank/DDBJ databases.
EMBL, AF36238; AAK545631; -
InterPro; IPR004109; HCV_NS3.
Pfam; PF02907; HCV_NS3; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Holland-Staley C.A., Kovari L.C., Golenberg E., Mayers D.L.;
"Genetic Diversity and response to IFN of the NS3 Protease Gene from Clinical Strains of the Hepatitis C Virus.";
Submitted (APR-2001) to the EMBL/GenBank/DDBJ databases.
EMBL; AF369241; AAK54566.1;
InterPro; IPR004109; HCV_NS3.
Pfam; PF02907; HCV_NS3; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Hepatitis C virus.
Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
Hepacivirus.
                                                                                                                                                                                                                                                  Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
Hepacivirus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Length 181;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     12; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            181 181 181 181 181 AM; 3B5E8161F2100A72 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                19123 MW; 1CAE817345ED809D CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               091RR2;
01-DEC-2001 (TrEMBLrel. 19, Created)
01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
NS3 protease (Fragment).
                                                                                                                                                                         Last sequence update)
Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Score 837; DB 12;
Pred. No. 8.5e-72;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  181 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     2; Mismatches
                                                                                                                                                     Created)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  PRT;
                                                                                                             PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        82.48;
                                                                                                                            091RRS;
01-DEC-2001 (TrEMBLrel. 19,
01-DEC-2001 (TrEMBLrel. 19,
01-MAR-2003 (TrEMBLrel. 23,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     164; Conservative
                                                                                                                                                                                                                 NS3 protease (Fragment). Hepatitis C virus.
                                                                                                             PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        181 1
181 AA;
                                                                                                                                                                                                                                                                                                                                           SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        NCBI_TaxID=11103;
                                                                                                                                                                                                                                                                                                 NCBI_TaxID=11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           STRAIN-Pt. 4V;
                                                                                                                                                                                                                                                                                                                                                               STRAIN-Pt.3U
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        NON_TER
NON_TER
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    NON_TER
NON_TER
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Protease.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Protease.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Local
                                                                                                          091RR5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Q91RR2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        RESULT 13
                                                                RESULT
Q91RR5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                091RR2
                                                                                                                              DATO DE LA PERENTA DE LA PEREN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ö
                                                                                                                                                                                                                                                                                                                                                                 ô
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       79 IASPKGPVTQMYTNVDKDLVGWQAPQGSRSLTPCTCGSSDLYLVTRHADVIPVRRRGDSR 138
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              79 IASPKGPVTQMYTNVDKDLVGWQAPQGSRSLTPCTCGSSDLYLVTRHADVIPVRRRGDSR 138
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          78
                                                                                                                                                                                                                                                                                                                                                                                                                             4 TAYAQQTRGLLGCIITSLTGRDKNQVEGEVQIVSTAAQTFLATCINGVCWTVYHGAGTRI 63
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               19 TAYAQQTRGEQGTQKTSHTGRDKNQVEGEVQIVSTATQTFLATSINGVLWTVYBGAGTRT 78
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      GSLLSPRPISYLKGSSGCPLLCPAGHAVGIFRAAVSTRGVAKAVDFIPVESLETTMRS 196
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            GSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVSTRGVAKAVDFIPVESLETTMRS 196
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                19 TAYAQQTRGEQGTQKTSHTGRDKNQVEGEVQIVSTATQTFLATSINGVLWIVYHGAGTRT
                                                     "Genetic Diversity and response to IFN of the NS3 Proteage Gene from Clinical Strains of the Hepatitis C Virus.";
Submitted (APR-2001) to the EMBL/GenBank/DDBJ databases.
EMBL, AF864205; ARK646601; -
InterPro; IPR004109; HCV_NS3.
Pfam, PP02907; HCV_NS3; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Holland-Staley C.A., Kovari L.C., Golenberg E., Mayers D.L.;
"Genetic Diversity and response to Irpn of the NS3 Protease Gene from
"Clinical Strains of the Hepatitis C Virus.";
Submitted (APR-2001) to the EMBL/GenBank/DDBJ databases.
EMBL, ARF86218, ARK5643.1: -.
InterPro: IPR004109; HCV_NS3.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Gaps
                                                                                                                                                                                                                                                                                                                                                                 Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
Hepacivirus.
                                                                                                                                                                                                                                                                                                                                                               ;
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          82.6%; Score 839; DB 12; Length 181; 92.7%; Pred. No. 5.5e-72;
                                                                                                                                                                                                                                                                                                                       Length 181;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       12; Indels
                                                                                                                                                                                                                                                                                                                                                               12; Indels
                                                                                                                                                                                                                                                         181 181
181 AA: 19130 MW; 85D91869299B7C35 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                181 AA; 19130 MW; 85D91869299B7C35 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Last sequence update)
Last annotation update)
                                                                                                                                                                                                                                                                                                                   Score 839; DB 12;
Pred. No. 5.5e-72;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             181 AA.
                                                                                                                                                                                                                                                                                                                                                             1; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Created)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          PRT;
                                                                                                                                                                                                                                                                                                                     82.68;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                01-DEC-2001 (TrEMBLrel. 19, 01-DEC-2001 (TrEMBLrel. 19, 01-MAR-2003 (TrEMBLrel. 23, NS3 protease (Fragment).
                                                                                                                                                                                                                                                                                                                                                             Matches 165; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Matches 165; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              181
                                                                                                                                                                                                                                                                                                                                         Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   NCBI_TaxID=11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    STRAIN-Pt. 4;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Protease.
                                                                                                                                                                                                                                   NON_TER
NON_TER
SEQUENCE
                                                                                                                                                                                                                   Protease.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          139
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      139
                                                                                                                                                                                                                                                                                                                     Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Local
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Q91RT5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                RESULT 11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ద
                                                                                                                                                                                                                                                                                                                                                                                                                                             윱
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       윰
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
```

Gaps

```
NCBI_TaxID-11103;
                                                                                                                                                                                                                                                                                           NON_TER
NON_TER
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                          Ouery Match
                                                                                                                                                                                                                                                                                                                                                                                                                   Best Loca
Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Search con
Job time
     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  음
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ö
                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            79 IASPKGPVIQMYINVDKDLVGWQAPQGSRSLIPCICGSSDLYLVIRHADVIPVRRKGDSR 138
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     64 IASPKGPVIQMYINVDKDLVGWPAPQGARSLTPCTCGSSDLYLVTRHADVIPVRRRGDSR 123
                                                                                                                                                                                                                 78
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         4 TAYAQQTRGLLGCIITSLTGRDKNQVBGEVQIVSTATQTFLATCINGVCWTVYHGAGTRT 63
                                                                                                                          78
                                                                                                                                                 4 TAYAQQIRGLLGCIXISLIGRDKNQVEGEVQIVSTAXQIFLATCINGVCWTVYHGAGIRI 63
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             GSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVSTRGVAKAVDFIPVESLETTMRS 196
                                                                                                                                                                                                                                                                                                                19 TAYAQQTRGEQGTQKTSHTGRDKNQVEGEVQIVSTATQTFLATSINGVLWTVYHGAGTRT
                                                                                                                        19 TAYAQQTRGEQGTQKTSHTGRDKNQVEGEVQIVSTATQTFLATSINGVLWTVYHGAGTRT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          STRAIN-PL.IC;
Holland-Staley C.A., Kovari L.C., Golenberg E., Mayers D.L.;
Holland-Staley C.A., Kovari L.C., Golenberg E., Mayers D.L.;
"Genetic Diversity and response to IFN of the NS3 Protease Gene from
Clinical Strains of the Hepatitis C Virus.";
Submitted (APR-2001) to the EMBL/GenBank/DDBJ databases.

EMBL, AF869214; AAK54539.1;
PinterPro; IPR004109; HCV_NS3.
Protease.
                                                                          Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Q91RR3;
01-DEC-2001 (TrEMBLrel. 19, Created)
01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
NS3 protease (Fragment).
NS3 protease (Fragment).
Hepatitis C virus.
Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            NG3 protease (Fragment).

Hepatitis C virus.
Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae; Hepacivirus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  .;
                                                                          .,
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Score 837; DB 12; Length 181;
Pred. No. 8.5e-72;
2; Mismatches 12; Indels (
                      Score 837; DB 12; Length 181; Pred. No. 8.5e-72;
                                                                       Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            181 181
181 AA; 19114 HW; BEIDOB542F014E86 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    01-DEC-2001 (TrEMBLrel. 19, Created)
01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
82.4%; Scur.
92.1%; Pred. No. o..
2; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                181 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        181 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 82.48;
92.18;
                                                                       Matches 164; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Matches 164; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Local Similarity
                                                Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          NCBI_TaxID-11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Hepaclvirus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            NON_TER
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          139
                         Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              091RR3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        091RT9
                                                                                                                                                                                                                                                                                                                                                                                                                                        RESULT 14
Q91RT9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                RESULT 15
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    091RR3
1D 092
1D 093
1D 093
1D 094
1D 095
1D
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                               셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ö
                                                                                                                                                           g
                                                                                                                                                                                                                    ŏ
                                                                                                                                                                                                                                                                                                                                                           a
                                                                                                                                                                                                                                                                                                                  ò
```

```
ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  79 IASPRGPVTQMYTNVDKDLVGWQAPQGSRSLTPCTCGSSDLYLVTRHADVIPVRRRGDSR 138
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           64 IASPKGPVIQMYINVDRDLVGWPAPQGARSLTPCTGGSSDLYLVTRHADVIPVRRRGDSR 123
                                                                                                                                                                                                                                                                                                                                                                                                                             78
                                                                                                                                                                                                                                                                                                                                                                                                                                                   4 TAYAQQTRGLLGCIITSLTGRDKNQVBGEVQIVSTAAQTFLATCINGVCWTVYHGAGTRT 63
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          139 GSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVSTRGVAKAVDFIPVESLETTMRS 196
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        "Genetic Diversity and response to IFN of the NS3 Protease Gene from Clinical Strains of the Hepatitis C Virus.";
Submitted (APR-2001) to the EMBL/Genbank/DDBJ databases.
EMBL; AF369240; AAK54565.1;
Interpro: IPR004109; HCV_NS3.
Pfam; PF02907; HCV_NS3; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                        19 TAYAQQTRGEQGTQKISHIGRDKNQVEGEVQIVSTATQTFLATSINGVLWTVYHGAGTRT
                                                                                                                                                                                                                                                                                                                                                                              Gaps
                                                                                                                                                                                                                                                                                                                                                                              ö
                                                                                                                                                                                                                                                                                                                                Length 181;
                                                                                                                                                                                                                                                                                                                                                                              Indels
                                                                                                                                                                                                                                                         181 181
181 AA; 19115 MW; 5D85F88AD7AC1A11 CRC64;
                                                                                                                                                                                                                                                                                                                           Score 836; DB 12;
Pred. No. 1.1e-71;
2; Mismatches 12;
                                                                                                                                                                                                                                                                                                                           tch 82.3%; al Similarity 92.1%; 164: Conservative
```

completed: August 30, 2003, 19:18:21 ne : 37.5921 secs

```
Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.
                                                                                                                                                                                                                                                                                                                                                                          SUMMARIES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ARI45273
ARI45271
ARI45271
ARI45255
ARI27810
BD081911
ARI27809
BD081910
ARI10828
BD081910
ARI10828
BD081910
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    AF009606
                                                                                                                                                                                                                                                    em_htgo_other:*
                                                             em_htg_other:*
em_htg_mus:*
                                                                                                                                                                                                        em_htgo_hum:*
em_htgo_mus:*
                                                                                                      em_htg_pln:*
em_htg_rod:*
em_htg_mam:*
                                                                                                                                                                   em_htg_vrt:*
                     em_htg_hum:*
                                                                                                                                                                                                                                                                                                                                                                                                                   Query
Match Length DB
em_v1:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     862.5
861.5
860.5
860.5
860.5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            857.5
857.5
856.5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  853.5
853.5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   RESULT 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Result
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Command line parameters:
-MODEL-frame+_pl. model -DEV-xlp
-MODEL-frame+_pl. model -DEV-xlp
-Gqn2_L/USPTO_spool/US09965594/runat_29082003_151919_28310/app_query.fasta_1.2872
-DB-GenEmbl -OFNT-fastap -SUPFIX-rge -MINMATCH-0.1 -LOOPCL-0 -LOOPEXT-0
-DB-GenEmbl -OFNT-fastap -SUPFIX-rge -MINMATCH-0.1 -LOOPCL-0 -LOOPEXT-0
-UNITS-blits -START-1 -END--1 -MATRIX-blosum62 -TRANS-human40 cdi -LIST-45
-DCALIGN-200 -THR_SCORE-pct -THR_MAX-100 -THR_MN-0 -ALIGN-15 -MODE-LOCAL
-USER-DCALIGN-200 -THR_SCORE-pct -THR_MAX-100 -MAXLEN-200000000
-USER-SCORE-SCORE-0 -WAIT -LIARGE GENE - ILL -SPBELOCK-100 -LONGLOG
-NO_MAAP -LARGE GUERY -NEG_SCORE-0 -WAIT -SPBELOCK-100 -LONGLOG
-DEV_TIMEOUT-120 -WARN_TIMEOUT-30 -THRADS-1 -XGAPDP-10 -XGAPEXT-0.5 -FCAPOP-6
-FCAPEXT-7 -YGAPDP-10 -YGAPEXT-0.5 -DELDP-6 -DELEXT-7
                                                                                                                                                                   (without alignments)
3147.423 Million cell updates/sec
                                                                                                                                                                                                                                               1016
1 MRKKGSVVIVGRINLSGDTA......VAKAVDFIPVESLETTMRSP 197
                                                                                                                                           August 30, 2003, 19:18:33; Search time 2560.57 Seconds
                 GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.
                                                                                                    - nucleic scarch, using frame_plus_p2n model
                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Total number of hits satisfying chosen parameters:
                                                                                                                                                                                                                                                                                                                                                                                                                                     2888711 seqs, 20454813386 residues
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries
                                                                                                                                                                                                                                                                                                                              0.5
                                                                                                                                                                                                                                                                                                                                . Xgapext
, Ygapext
, Fgapext
, Delext
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Minimum DB seq length: 0
Maximum DB seq length: 2000000000
                                                                                                                                                                                                                              US-09-965-594-22
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    9b_ba:*
9b_htg:*
9b_on:*
9b_on:*
9b_pat:*
9b_ph:*
9b_ph:*
9b_ph:*
9b_pr:*
9b_pr:*
9b_vi:*
em_ba:*
em_ba:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  em_ov:*
em_pat:*
em_ph:*
em_pl:*
em_ro:*
em_ro:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                em_hum:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    em_in:*
                                                                                                                                                                                                                                                                                                          BLOSUM62
Xgapop 10.0 ,
Ygapop 10.0 ,
Fgapop 6.0 ,
Delop 6.0 ,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           em_om: *
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 em_or:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     GenEmbl:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 22223325
2222333
222233
222233
22233
2223
2223
2223
2223
2223
2223
2223
2223
2223
2223
2223
2223
2223
2223
2223
2223
2223
2223
2223
2223
2223
2223
2223
2223
2223
2223
2223
2223
2223
2223
2223
2223
2223
2223
2223
2223
2223
2223
2223
2223
2223
2223
2223
2223
2223
2223
2223
2223
2223
2223
2223
2223
2223
2223
2223
2223
2223
2223
2223
2223
2223
2223
2223
2223
2223
2223
2223
2223
2223
2223
2223
2223
2223
2223
2223
2223
2223
2223
2223
2223
2223
2223
2223
2223
2223
2223
2223
2223
2223
2223
2223
2223
2223
2223
2223
2223
2223
2223
2223
2223
2223
2223
2223
2223
2223
2223
2223
2223
2223
2223
2223
2223
2223
2223
2223
2223
2223
2223
2223
2223
2223
2223
2223
2223
2223
2223
2223
2223
2223
2223
2223
2223
2223
2223
2223
2223
2223
2223
2223
2223
2223
2223
2223
2223
2223
2223
2223
2223
2223
2223
2223
2223
2223
2223
2223
2223
2223
2223
2223
2223
2223
2223
2223
2223
2223
2223
2223
2223
2223
2223
2223
2223
2223
2223
2223
2223
2223
2223
2223
2223
2223
2223
2223
2223
2223
2223
2223
2223
2223
2223
2223
2223
2223
2223
2223
2223
2223
2223
2223
2223
2223
2223
2223
2223
2223
2223
2223
2223
2223
2223
2223
2223
2223
2223
2223
2223
2223
2223
2223
2223
2223
2223
2223
2223
2223
2223
2223
2223
2223
2223
2223
2223
2223
2223
2223
2223
2223
2223
2223
2223
2223
2223
2223
2223
2223
2223
2223
2223
2223
2223
2223
2223
2223
2223
2223
2223
2223
2223
2223
2223
2223
2223
2223
2223
2223
2223
2223
2223
2223
2223
2223
2223
2223
2223
2223
2223
2223
2223
2223
2223
2223
2223
2223
2223
2223
2223
2223
2023
2023
2023
2023
2023
2023
2023
2023
2023
2023
2023
2023
2023
2023
2023
2023
2023
2023
2023
2023
2023
2023
2023
2023
2023
2023
2023
2023
2023
2023
2023
2023
2023
2023
2023
2023
2023
2023
2023
2023
2023
2023
2023
2023
2023
2023
2023
2023
2023
2023
2023
2023
2023
2023
2023
2023
2023
2023
2023
2023
2023
2023
2023
2023
2023
2023
2023
2023
2023
2023
2023
2023
2023
2023
2023
2023
2023
2023
2023
2023
2023
2023
2023
2023
2023
2023
2023
2023
2023
2023
2023
2023
2023
2023
2023
2023
2023
2023
2023
2023
2023
2023
2023
2023
2023
2023
2023
2023
2023
2023
2023
2023
2023
2023
2023
                                                                                                                                                                                                                                                  Perfect score:
                                                                                                                                                                                                                                                                                                               Scoring table:
                                                                                                      OM protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Database :
                                                                                                                                                                                                                                                                          Sequence:
                                                                                                                                                                                                                                                                                                                                                                                                                                       Searched:
                                                                                                                                             uo:
```

AR145253 Sequence AR145266 Sequence AR145261 Sequence AR145269 Sequence AR268278 Pestiviru AR145257 Sequence

Sequence

Sequence

Sequence

Description

Sequence

ALIGNMENTS

ARI10828 Sequence BD069982 Functiona AF009606 Hepatitis ARI10831 Sequence BD069985 Functiona

ARI18692 Sequence 106440 Sequence 54 109329 Sequence 10 ARI18696 Sequence

109331 Sequence 15 M32084 Hepatitis ( AR118703 Sequence

AR118686 Sequence IO6434 Sequence 48 IO9328 Sequence 8

AR145259 Sequence AR127810 Sequence BD081911 Hepatitis AR127809 Sequence BD081910 Hepatitis AR110828 Sequence

Sequence Sequence Sequence Sequence

Sequence Sequence Sequence Sequence

Sequence

Sequence

```
.1998
                                     Unknown.
Unclassified.
AR145264
AR145264.1
                                                                                                                                                                                  Percent Similarity:
Best Local Similarity:
Query Match:
                                                                                                                              ď
                           Unknown
                                                                                                                                                           Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                82
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            142
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       604
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       RESULT 3
AR145268
LOCUS
DEFINITION
ACCESSION
VERSION
                                                                                                                             BASE COUNT
ORIGIN
                                                                                                                                                                  Pred. No.:
                                    ORGANISM
                                                                AUTHORS
TITLE
ACCESSION
                                                                                         JOURNAL
        VERSION
KEYWORDS
                                                       REFERENCE
                                                                                                   FEATURES
                            SOURCE
                                                                                                                                                                           Score:
                                                                                                                                                                                                                                                                                                                                                                                                                 8
                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                  a
                                                                                                                                                                                                                                                                                                                                         qq
                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                       a
                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              PAT 08-AUG-2001
         PAT 20-APR-2002
                                                                                                                                                                                                                                                                                                                                                                                                                                                    ValThrArgHisAlaAspValIleProValArgArgArgGlyAspSerArgGlySerLeu 141
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              142 LeuSerProArgProlleSerTyrLeuLysGlySerSerGlyGlyProLeuLeuCysPro 161
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                892
                                                                                                                                                                                                                                                                                                                         473 GCCCAGCAGAGAGGCCTCCTAGGGTGTAAGATCACCAGTCTGACTGGCCGGGACAAA 532
                                                                                                                                                                                                                                                                                                                                                                                                                  ProLysGlyProValThrGlnMetTyrThrAsnValAspLysAspLeuValGlyTrpGln 101
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   162 AlaGlyHisAlaValGlyIlePheArgAlaAlaValSerThrArgGlyValAlaLysAla 181
                                                                                                                                                                                                                                                                                                        41
                                                                                                                                                                                                                                                                                                                                           61
                                                                                                                                                                                                                                                                                                                                                                               81
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        22 AlaGlnGlnThrArgGlyGluGlnGlyThrGlnLysThrSerHisThrGlyArgAspLys
                                                                                                                                                                                                                                                                                                                                           AsnGlnValGluGlyGluValGlnIleValSerThrAlaThrGlnThrPheLeuAlaThr
                                                                                                                                                                                                                                                                                                                                                                                SerIleAsnGlyValLeuTrpThrValTyrHisGlyAlaGlyThrArgThrIleAlaSer
                                                                                1 (bases 1 to 12734)
Hong.Z., Lai,W.C.H. and Lau,J.Y.N.
Hepatitis C virus protease-dependent chimeric pestivirus Patent: US 6326137-A 1 04-DEC-2001;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ValAspPheIleProValGluSerLeuGluThrThrMetArgSer 196
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               linear
         linear
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  12734
176
4
12
3
                                                                                                                                                                                   Length:
Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               DNA
         DNA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              AR145264 1998 bp D
Seguence 105 from patent US 6211338.
       AR179057 12734 bp
Sequence 1 from patent US 6326137
AR179057.1 GI:20220612
                                                                                                                                                                                                                                                  US-09-965-594-22 (1-197) x AR179057 (1-12734)
                                                                                                                     Location/Qualiflers
                                                                                                                                      /organism-"unknown"
2604 c 3295 g
                                                                                                                                                                                  3.77e-64
882.50
92.31%
90.26%
                                                                        Unclassified
                                                                                                                                                                                                              Best Local Similarity:
Query Match:
DB:
                                                                                                                                                 4032 a
                                                      Unknown.
                                                                                                                                                                                              Score:
Percent Similarity:
                                                                 Unknown
                                                                                                                                                                            Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      893
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        182
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         953
                                                                                                                                                                                                                                                                                                                                                              533
                                                                                                                                                                                                                                                                                                                                                                               62
                                                                                                                                                                                                                                                                                                                                                                                                                   82
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          122
                                                                                                                                                                                                                                                                                                                                            42
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        DEFINITION
       LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
                                                                                                                                source
                                                                                                                                                BASE COUNT
                                                                                                                                                                                      Pred. No.:
                                                                                                   TITLE
JOURNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            RESULT 2
AR145264
LOCUS
                                                                                           AUTHORS
                                                                                 REFERENCE
                                                                                                                     FEATURES
AR179057
                                                                                                                                                           ORIGIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ద
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                      q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                        ð
                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                         õ
                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                  셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           6
```

```
08-AUG-2001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  243
                                                                                                                                                                                                                                                                                                                                                                                                                                                                    161
                                                                                                                                                                                                                                                                                                                     101
                                                                                                                                                                                                                                                                                                                                                                                   102 AlaProGlnGlySerArgSerLeuThrProCysThrCysGlySerSerAspLeuTyrLeu 121
                                                                                                                                                                                                                                                                                                                                                                      41
                                                                                                                                                                                                                                                                                                                                                                                                                     61
                                                                                                                                                                                                                                                                                                                                                                                                                    ProLysGlyProValThrGlnMetTyrThrAsnValAspLysAspLeuValGlyTrpGln
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    LeuSerProArgProlleSerTyrLeuLysGlySerSerGlyGlyProLeuLeuCysPro
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             484 CICICCCCAGGCCIGICICCTACITGAAGGCICTICGGGIGGICCICCACIGCICICCC
                                                                                                                                                                                                                                                                                                                                                                      22 AlaGlnGlnThrArgGlyGluGlnGlyThrGlnLysThrSerHisThrGlyArgAspLys
                                                                                   NS3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                PAT
                                                          1 (bases 1 to 1998)
Malcolm, B.A., Taremi.S.Shane., Weber, P.C. and Yao, N.
Single-chain recombinant complexes of hepatitis C virus
protease and MS4A cofactor peptide
Patent: US 6211338-A 105 03-APR-2001;
Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      182 ValaspPheIleProValGluSerLeuGluThrThrMetArgSerPro 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  linear
                                                                                                                                                                                                        1998
167
15
11
                                                                                                                                                                                                           Length:
Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  DNA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               AR145268 1998 bp D
Sequence 109 from patent US 6211338.
AR145268.1 GI:15107135
                                                                                                                                                            423
                                                                                                                                                                                                                                                                                                US-09-965-594-22 (1-197) x AR145264 (1-1998)
                                                                                                                                              /organism="unknown"
595 c 569 g
GI:15107131
                                                                                                                                                                                                          1.73e-64
875.50
92.86%
85.20%
86.17%
```

```
Unknown.
                                                                                                        Alignment Scores
                                                                                                                                                                                                                                                       42
                                                                                                                                                                                                                                                                                                                      82
                                                                                                                                                                                                                                                                                                                                                                      364
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     604
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             AR145254
LOCUS
DEFINITION
ACCESSION
VERSION
VERSION
SOURCE
ORGANISM
                                                                                                                                               Query Match:
                                                                 source
ORGANISM
                                                                               BASE COUNT
                                                                                                                ..
&
                                                JOURNAL
                REFERENCE
                         AUTHORS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       RESULT 5
                                                         FEATURES
                               TITLE
                                                                                         ORIGIN
                                                                                                                                                                                                                                                                                                                                                                                                     a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     음
                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                      à
                                                                                                                                                                                                                                                                  qq
                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                 δ
                                                                                                                                                                                                                                                                                                                                  ద
                                                                                                                                                                                                                                                                                                                                                      õ
                                                                                                                                                                                                                                                                                                                                                                   a
                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                      ö
                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                    ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                       ò
                                                                                                                                                                                                      a
                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                     08-AUG-2001
                                                                                                                                                                                                                                                                                       243
                                                                                                                                                                                                                                                                                                             244 TGCGTCAACGGCGTGTGTTGGACCGTTTACCATGGTGCTGGCTCAAAGACCTTAGCCGGC 303
                                                                                                                                                                                                                                                                                                                                                                     122 ValThrArgHisAlaAspVallleProValArgArgArgGlyAspSerArgGlySerLeu 141
                                                                                                                                                                                                                                                                                                                                                                                                              LeuSerProArgProIleSerTyrLeuLysGlySerSerGlyGlyProLeuLeuCysPro 161
                                                                                                                                                                                                                                                                                                                                                                                                                                                                     AlaciyhisAlaValGiyIlePheArgAlaAlaValSerThrArgGiyValAlaLysAla 181
                                                                                                                                                                                                     4.
                                                                                                                                                                                                                                                                                                     62 SerileAsnGlyValLeuTrpThrValTyrHisGlyAlaGlyThrArgThrIleAlaSer 81
                                                                                                                                                                                                                                                                       42 AsnGlnvalGluGlyGluValGlnIleValSerThrAlaThrGlnThrPheLeuAlaThr 61
                                                                                                                                                                                                                                                                                                                                                                                                                                                AlaGlnGlnThrArgGlyGluGlnGlyThrGlnLysThrSerHisThrGlyArgAspLys
                                                NS3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       PAT
                              Malcolm, B.A., Taremi, S.Shane., Weber, P.C. and Yao, N. Single-chain recombinant complexes of hepatitis C virus protease and NS4A cofactor peptide Patent: US 6211338-A 109 03-APR-2001; Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     linear
                                                                                                                              1998
166
16
11
3
                                                                                                                             Length:
Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        DNA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1998 bp D
Sequence 104 from patent US 6211338.
AR145263
                                                                                                                                                                                      (1-1998)
                                                                                       /organism-"unknown"
595 c 569 g
                                                                                                                                                                                       US-09-965-594-22 (1-197) x AR145268
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               AR145263.1 GI:15107130
                                                                                                                              3.09e-64
872.50
92.86%
84.69%
85.88%
                                                                                 .1998
                        Jnclassified
                                                                                                                                           Percent Similarity:
Best Local Similarity:
Query Match:
DB:
                                                                                                 a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Unknown
        Unknown
                 Unknown
                                                                                                                        Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                     184
                                                                                                                                                                                                                                                                                                                                                                                                                                                                      162
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      182
                                                                                                                                                                                                                                                                                                                                                                                                                                      142
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      604
                                                                                                                                                                                                                                       22
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
                                                                                              BASE COUNT
ORIGIN
               ORGANISM
                                                                                                                                ..
Q
                                        AUTHORS
TITLE
                                                                JOURNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      RESULT 4
AR145263
LOCUS
                                REFERENCE
KEYWORDS
                                                                        FEATURES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      용
                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                    ογ
                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ô
                                                                                                                                                                                                                                                                                  a
                                                                                                                                                                                                      ò
                                                                                                                                                                                                                     qq
                                                                                                                                                                                                                                     Q
                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                     õ
                                                                                                                                                                                                                                                                                                     ô
                                                                                                                                                                                                                                                                                                                  셤
                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                   음
                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                  q
                                                                                                                                                                                                                                                                                                                                                                                                     8
```

```
PAT 08-AUG-2001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    161
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          543
                                                                                                                                                                                                                                                     183
                                                                                                                                                                                                                                                                                                                                                                      243
                                                                                                                                                                                                                                                                                                                                                                                             101
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        102 AlaProGlnGlySerArgSerLeuThrProCysThrCysGlySerSerAspLeuTyrLeu 121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          61
                                                                                                                                                                                                                                                                                                    41
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 AsnGlnValGluGlyGluValGlnIleValSerThrAlaThrGlnThrPheLeuAlaThr
                                                                                                                                                                                                                                                                                                                                                            ProLysGlyProValThrGlnMetTyrThrAsnValAspLysAspLeuValGlyTrpGln
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               162 AlaGlyHisAlaValGlyIlePheArgAlaAlaValSerThrArgGlyValAlaLysAla
                                                                                                                                                                                                                                                                                                    22 AlaGlnGlnThrArgGlyGluGlnGlyThrGlnLysThrSerHisThrGlyArgAspLys
                                  NS3
        1 (bases 1 to 1998)
Malcolm,B.A., Taremi,S.Shane., Weber,P.C. and Yao,N.
Single-chain recombinant complexes of hepatitis C virus
protease and NSA cofactor peptide
Patent: US 6211338-A 104 03-APR-2001;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             182 ValAspPheIleProValGluSerLeuGluThrThrMetArgSerPro 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   651
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         linear
                                                                                                                                                 1998
167
14
12
3
                                                                                                                                                Length:
Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   NA
NA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                651 bp
from patent US 6211338.
                                                                                                    424
                                                                                                                                                                                                           Gaps:
                                                                                                                                                                                                                               US-09-965-594-22 (1-197) x AR145263 (1-1998)
                                                                Location/Qualifiers
                                                                                         /organism="unknown"
596 c 568 g
                                                                                                 568 9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    GI:15107121
                                                                                                                                                3.75e·64
871.50
92.35%
85.20%
85.78%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  AR145254
Sequence 95 1
AR145254
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Unknown.
Unclassified.
Unclassified
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    AR145254.1
                                                                                                                                                                       Percent Similarity:
Best Local Similarity:
```

of hepatitis

```
Unknown
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Unknown
                                                                                 Scores:
                                                                                                                                                                        'n
                                                                                                                                                                                                                                                                                                                                                                                    122
                                                                                                                                                                                                                                                                                                                                                                                                                        142
                                                                                                                                                                                                                                                                                                                                                                                                                                       484
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             182
                                                                                                                                                                                                                                                                                                               82
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        DEFINITION
ACCESSION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   KEYWORDS
SOURCE
ORGANISM
                                                      BASE COUNT
ORIGIN
                                                                                        Pred. No.:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      REFERENCE
AUTHORS
TITLE
                                                                                 Alignment
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               RESULT 7
AR145262
LOCUS
                   JOURNAL
                            FEATURES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           VERSION
  TITLE
                                                                                                  Score:
                                                                                                                                                                        δ
                                                                                                                                                                                       a
                                                                                                                                                                                                         ò
                                                                                                                                                                                                                         엄
                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                            ö
                                                                                                                                                                                                                                                                                           a
                                                                                                                                                                                                                                                                                                               ö
                                                                                                                                                                                                                                                                                                                               qq
                                                                                                                                                                                                                                                                                                                                                  õ
                                                                                                                                                                                                                                                                                                                                                                qq
                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                    q
                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                      qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                         õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                         5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  PAT 08-AUG-2001
                                                                                                                                                                                                                                                                                                                              303
                                                                                                                                                                                                                                                                                                                                                                                                    ValThrArgHisAlaAspValIleProValArgArgArgGlyAspSerArgGlySerLeu 141
                                                                                                                                                                                                                                                                                                                                                                                                             543
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             603
                                                                                                                                                                                                 64 GGTTCTGTTGTTGTTGCTAGAATTATTTTATCTGGTAGTGCTAGTATCACGGCCTAC 123
                                                                                                                                                                                                                                  LeuSerProArgProlleSerTyrLcuLysGlySerSerGlyGlyProLeuLeuCysPro 161
                                                                                                                                                                                                                                                                                                                                                                                                                                                                          162 AlaGlyHisAlaValGlyIlePheArgAlaAlaValSerThrArgGlyValAlaLysAla 181
                                                                                                                                                                                                                         41
                                                                                                                                                                                                                                                           61
                                                                                                                                                                                                                                                                                            81
                                                                                                                                                                                        -----ThrAlaTyr 21
                                                                                                                                                                                                                                                                                            SerileasnGlyValLeuTrpThrValTyrHisGlyalaGlyThrArgThr1leAlaSer
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   22 AlaGlnGlnThrArgGlyGluGlnGlyThrGlnLysThrSerHisThrGlyArgAspLys
                                                                                                                                                                                                                                                           AsnGlnValGluGlyGluValGlnIleValSerThrAlaThrGlnThrPheLeuAlaThr
1 (bases 1 to 651).
Malcolm,B.A., Taremi,S.Shane., Weber,P.C. and Yao,N.
Single-chain recombinant complexes of hepatitis C virus
protease and NG4A cofactor peptide
Patent: US 6211338-A 95 03-APR-2001;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1 (bases 1 to 1998)
Malcolm, B.A., Taremi, S.Shane., Weber, P.C. and Yao, N.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ValAspPheIleProValGluSerLeuGluThrThrMetArgSer 196
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   linear
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       GlyServalValIleValGlyArgIleAsnLeuSerGlyAsp---
                                                                                                         651
166
15
11
11
                                                                                                        Length:
Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   DNA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Sequence 108 from patent US 6211338.
AR145267
AR145267.1 G1:15107134
                                                                       144
                                                                                                                                                     Gaps:
                                                                                                                                                                      US-09-965-594-22 (1-197) x AR145254 (1-651)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1998
                                            Location/Qualifiers
                                                              /organism="unknown"
                                                                      200 9
                                                                                                        1.84e-64
868.50
92.82%
85.13%
                                                                      187 c
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Unknown.
Unclassified.
                                                                                                                          Percent Similarity:
Best Local Similarity:
                                                                      120 a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  AR145267
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Unknown.
                                                                                                  Alignment Scores:
                                                                                                                                                                                                                                                           42
                                                                                                                                                                                                                                                                                            62
                                                                                                                                                                                                                                                                                                               244
                                                                                                                                                                                                                                                                                                                                82
                                                                                                                                                                                                                                                                                                                                                 304
                                                                                                                                                                                                                                                                                                                                                                                                     122
                                                                                                                                                                                                                                                                                                                                                                                                                                       142
                                                                                                                                                                                                                                                                                                                                                                                                                                                        484
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             182
                                                                                                                                            Query Match:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        DEFINITION
ACCESSION
VERSION
                                                       source
                                                                      BASE COUNT
ORIGIN
                                                                                                          Pred. No.:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ORGANISM
REFERENCE
AUTHORS
TITLE
                                     JOURNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               RESULT 6
AR145267
LOCUS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    KEYWORDS
SOURCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        REFERENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                AUTHORS
                                             FEATURES
                                                                                                                    Score:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                       a
                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                            qq
                                                                                                                                                                                                                                                                                                                                                                                                                    8
                                                                                                                                                                                                                                                                                                                                                                                                                                                       d
                                                                                                                                                                                                                                                                                                                                                                                                                                                                       δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            δ
                                                                                                                                                                                       ò
                                                                                                                                                                                                       a
                                                                                                                                                                                                                         à
                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                       ö
                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                             셤
                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                  a
                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                      õ
```

```
PAT 08-AUG-2001
                                                                                                                                                                                                                                                                                                    141
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               123
                                                                                                                                                                                                                                                                                                                                                                                                                                                         ProLysGlyProValThrGlnMetTyrThrAsnValAspLysAspLeuValGlyTrpGln 101
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     304 CCAAAGGGGCCAATCACCCAGATGTACACTAATGTGGACCAGGACCTCGTCGGCTGGCAG 363
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    LeuSerProArgProlleSerTyrLeuLysGlySerSerGlyGlyProLeuLeuCysPro 161
                                                                                                                                                                                                                                                                                                                                                        61
                                                                                                                                                                                                                                                                                                                                                                                                        81
                                                                                                                                                                                                                                                    7
                                                                                                                                                                                                                                                                   ValThrArgHisAlaAspValIleProValArgArgArgGlyAspSerArgGlySerLeu
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   GlySerValValIleValGlyArgIleAsnLeuSerGlyAsp-----ThrAlaTyr
                                                                                                                                                                                                                                                                                                                                                       42 AsnGlnValGluGlyGluValGlnIleValSerThrAlaThrGlnThrPheLeuAlaThr
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           (bases 1 to 1998)
Malcolm, B.A., Taremi, S. Shane., Weber, P.C. and Yao, N.
Single-chain recombinant complexes of hepatitis C virus
protease and NS4A cofactor peptide
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      linear
                                                                                                                               1998
166
15
12
3
                                                                                                                               Length:
Matches:
Conservative:
Mismatches:
Indels:
Single-chain recombinant complexes of
protease and NS4A cofactor peptide
Patent: US 6211338-A 108 03-APR-2001;
Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Sequence 103 from patent US 6211338. AR145262
                                                                               424
                                                                                                                                                                                                   Gaps:
                                                                                                                                                                                                                          (1-1998)
                                                                /organism="unknown"
596 c 568 q
                                                                                                                                                                                                                          US-09-965-594-22 (1-197) x AR145267
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            AR145262.1 GI:15107129
                                                                                                                               6.72e-64
868.50
92.35%
84.69%
85.48%
                                                     1. .1998
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Unclassified
                                                                                                                                                         Percent Similarity:
Best Local Similarity:
Query Match:
DB:
                                                                               Ø
```

irce 1651 /organism="unknown" iNT 120 a 187 c 200 g 144 t it Scores: 3.29e-64 Length: 0.: 865.50 Matches:	February   10   February   10   February   10   February   11   February   11   February   12   February   13   February   14   February   14   February   15   February   16   February   1	Oy 5 GlySerValValIleValGlyArglleAsnLeuSerGlyAspThralaTyr 21	Db 64 GGTTCTGTTATTGTTAGGTAGAATTATTTTATCTGGTAGTAGTATCACGGCCTAC 123	Qy         22 AladloGlnThrArgGlyGluGlnGlyThrGlnLysThrSerHisThrGlyArgAspLys 41           :::	42 AsnGlnValGluGlyGluValGlnIleValSerThrAlaThrGlnThrPheLeuAlaThr	Db 184 AACCAGGTCGAGGAGAGGTTCAGGTGGTTTCCACGCAACAACATCTTCCTGGCGACC 243  Qy 62 SerlleAsnGlyValLeuTrpThrValTyrHisGlyAlaGlyThrArqThrIleAlaSer 81	244	82 ProLysGlyProValThrGlnMetTyrThrAsnValAspLysAspLeuValGlyTrpGln	Up 102 AlaProGlnGlySerArgSerLeuThrProCysThrCysGlySerSerAspLeuTyrLeu 121		122 ValThrargHisalaAspVallleProValArgArgArgGlyAspSerArgGlySerLeu 14	DD 424 GTCACGAGACATGCTGACGTCATTCCGGTGCGCGGGGGGGG	Db 484 CTCTCCCCAGGCCTGTCTCTACTTGAAGGCTCTGCTGGTGGTCCACTCTCTGCCCT 543	Qy 162 AlaGlyHisAlaValGlyIlePheArgalaAlaValSerThrArgGlyValAlaLysAla 181 :::	Oy 182 ValAspPhelleProValGluSerLeuGluThrThrMetArgSer 196	604		ISM UNKNOWN. Unclassified. CE 1 (bases 1 to 651) RS Malcolm.B.A., Taremi,S.Shane., Weber,P.C. and Yao,	TITLE Single-chain recombinant complexes of hepatitis C virus NS3 protease and NS4A cofactor peptide JOURNAL Patent: US 6211338-A 94 03APR-2001; FEATURES Location/Qualiflers   1.651   1.651     1.651   1.651     1.651   1.651     1.651   1.651     1.651   1.651     1.651   1.651     1.651   1.651     1.651   1.651     1.651   1.651     1.651   1.651     1.651   1.651     1.651   1.651     1.651   1.651     1.651   1.651     1.651   1.651     1.651   1.651   1.651     1.651   1.651   1.651     1.651   1.651   1.651     1.651   1.651   1.651   1.651     1.651   1.651   1.651   1.651     1.652   1.652   1.652   1.652   1.652     1.653   1.652   1.652   1.652   1.652   1.652     1.653   1.652   1.652   1.652   1.652   1.652     1.653   1.652   1.652   1.652   1.652   1.652     1.654   1.652   1.652   1.652   1.652   1.652     1.655   1.652   1.652   1.652   1.652   1.652   1.652     1.655   1.652   1.652   1.652   1.652   1.652   1.652     1.655   1.652   1.652   1.652   1.652   1.652   1.652     1.655   1.652   1.652   1.652   1.652   1.652   1.652     1.655   1.652   1.652   1.652   1.652   1.652   1.652     1.655   1.652   1.652   1.652   1.652   1.652   1.652     1.655   1.652   1.652   1.652   1.652   1.652   1.652   1.652   1.652     1.655   1.652   1.65
103 03-APR-2001; llifiers inknown* 568 g 424 t	• 6) 6) >	US-09-965-594-22 (1-197) x AR145262 (1-1998)	Oy 5 GlySerValValIleValGlyArglleAsnLeuSerGlyAspThrAlafyr 21	Db 64 GGTTCTGTTGTTGTTGGTAGATTATTTATCTGGTAGTAGTATCACGGCCTAC 123 Ov 22 AlaGlnGlnThrArgGlvGluGlnThrGlnLvsThrSerHisThrGlvArgAspLvs 41	:::	Qy 42 AsnGlnValGluGlyGluValGlnIleValSerThrAlaThrGlnThrPheLeuhlaThr 61	62 SerIleAsnGlyValLeuTrpThrValTyrHisGlyAlaGlyThrArgThrIleAlaSer	244	CCAAAGGGCCAATCACCAGATGTACATGTGGACCAGGACCTGGTGGCGGCGGGGCGGGGGGGG	Oy 102 AlabroGinGlySerArgSerLeuThrProCysThrCysGlySerSerAspLeuTyrLeu 121	364 GCGCCCCCGGGGCGCTTCCTTGACACCATCCTGTGGCACCTCTATACTTG	UY 12. WALTHIAGHISAARSPVALLEPROVALAGRGANGULYASDFERGINGULYEL 141	Oy 142 LeuSerProArgProlleSerTyrLeuLysGlySerSerGlyGlyProLeuLeuCysPro 161	Db 484 CTCTCCCCCAGGCCTGTCTCCTACTTGAAGGCTCTTCGGGTGGTCCACTGCTCTGCCCT 543  Oy 162 AlaGlyHisAlaValGlyIlePheArgAlaAlaValSerThrArgGlyValAlaLySAla 181		<pre>Qy 182 ValAspPheIleProValGluSerLeuGluThrThrMetArgSerPro 197                                     </pre>	RESULT 8 AR145258 LCCUS AR145258 AR145258 DEFINITION Sequence 99 from patent US 6211338. ACESSION AR14228 AVENSION AR145258.1 GI:15107125	×	NEFERENCE 1 (Dases 1 to 51)  AUTHORS Malcolm, B.A., Tartemi.S.Shane., Weber, P.C. and Yao, N.  TITLE Single-chain recombinant complexes of hepatitis C virus NS3  TOTRE protease and NS4A cofactor peptide  JOHNNAL Patent: US 6211338-A 99 03-APR-2001;  FEATURES Location/Qualifiers

```
Alignment Scores
                                                                                        ð
                                                                                                                                                                                                                                                                                                                                                                                                                                              PAT 08-AUG-2001
                                                                                                                                                                                                                                  243
                                                                                                                                                                                                                                                                  102 AlaProGlnGlySerArgSerLeuThrProCysThrCysGlySerSerAspLeuTyrLeu 121
                                                                                                                                                                     61
                                                                                                                                     41
                                                                                                      21
                                                                                                                                                                   22 AlaGlnGlnThrArgGlyGluGlnGlyThrGlnLysThrSerHisThrGlyArgAspLys
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Unclassified.

[ (bases 1 to 1998)]

Malcolm,B.A., Taremi,S.Shane., Weber,P.C. and Yao,N.

Sigle-chain recombinant complexes of hepatitis C virus NS3
protease and NSA cofactor peptide
Patent: US 6211338-A 107 03-APR-2001;
                                                                                                                                                                                                                                                                                                                                                                                                               648
                                                                                                                                                                                                                                                                                                                                                                                               182 ValAspPhelleProValGluSerLeuGluThrThrMetArgSer 196
                                                                                                                                                                                                                                                                                                                                                                                                                                             linear
                                                                                                                                                                                                                                                                                                                                                                                                        5 GlySerValValIleValGlyArgIleAsnLeuSerGlyAsp--
                                651
166
14
13
                              Length:
Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ų
                                                                                                                                                                                                                                                                                                                                                                                                                                                    Sequence 107 from patent 05 6211338.
AR145266
AR145266.1 GI:15107133
145
                                                                      Gaps:
                                                                                                                                                                                                                                                                                                                                                                                                                                              þþ
                                                                                      US-09-965-594-22 (1-197) x AR145253 (1-651)
                                                                                                                                                                                                                                                                                                                                                                                                                                             1998
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     /organism="unknown"
596 c 568 q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              568 9
b
199
                              4e-64
864.50
92.31%
85.13%
85.09%
U
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1998
                                          Percent Similarity:
Best Local Similarity:
Query Match:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ø
                                                                                                                                                                                                                                                                                                                                                                                                                                             AR145266
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Unknown.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Unknown.
                        Alignment Scores:
                                                                                                                                                                                                                                                                                 364
                                                                                                                                                                                                                                                                                                                                                                                                                                                   DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            BASE COUNT
ORIGIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ORGANISM
                                ..
9
                                                                                                                                                                                                                                                                                                                                                                                                                             RESULT 10
AR145266
LOCUS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             REFERENCE
AUTHORS
BASE COUNT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             JOURNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       FEATURES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             TITLE
      DRIGIN
                                Pred.
                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                            q
                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                            õ
                                                                                                                                                                                                                                                                                                                                                                                                            q
                                                                                                      ð
                                                                                                                  g
                                                                                                                                     ò
                                                                                                                                                 ద
                                                                                                                                                                   ò
                                                                                                                                                                                  d
                                                                                                                                                                                                   à
                                                                                                                                                                                                               g
                                                                                                                                                                                                                                  ð
                                                                                                                                                                                                                                                 셤
                                                                                                                                                                                                                                                              οχ
                                                                                                                                                                                                                                                                                                 õ
                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                             셤
                                                                                                                                                                                                                                                                                                                                                               ä
```

```
08-AUG-2001
                                                                                                                                                                                                                                                                                  121
                                                                                                                                                                                                                                                                    423
                                                                                                                                                                                                                                                                                                                      161
                                                                                                                                                                                                                                                                                                                                                              243
                                                                                                                                                                                                    303
                                                                                                                                                                                                                    101
                                                                                                                                                                                                                                                                                                                                                     162 AlaGlyHisAlaValGlyIlePheArgAlaAlaValSerThrArgGlyValAlaLySAla 181
                                                                                          4.1
                                                                                                                                                    61
                                                                                                                                                                                    81
                                                                                                                                                                                                                                                                                                                   244 TGCGTCAACGGCGTGTGGACCGTTACCATGGTGCTGGCTCAAAGACCTTAGCCGGC
                                                                                                                ProLysGlyProValThrGlnMetTyrThrAsnValAspLysAspLeuValGlyTrpGln
                                                                                  GlySerValValIleValGlyArglleAsnLeuSerGlyAsp.----ThrAlaTyr
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1 (bases 1 to 1998)
Malcolm, B.A., Taremi, S.Shane, Weber, P.C. and Yao, N.
Single-chain recombinant complexes of hepatitis C virus NS3
protease and N84A cofactor peptide
Patent: US 6211338-A 102 03-APR-2001;
                                                                                                                                                                                                                                                                                                                                                                                                                                       PAT
                                                                                                                                                                                                                                                                                                                                                                                      ValAspPhelleProValGluSerLeuGluThrThrMetArgSerPro 197
                                                                                                                                                                                                                                                                                                                                                                                                      651
                                                                                                                                                                                                                                                                                                                                                                                                linear
       1998
165
16
12
3
               Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                    Sequence 102 from patent US 6211338. AR145261 GI:15107128
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           425
                                                                  US-09-965-594-22 (1-197) x AR145266 (1-1998)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Location/Qualifiers
1. .1998
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 /organism="unknown"
597 c 567 g
       1.46e-63
864.50
92.35%
84.18%
85.09%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Unclassified
                        Percent Similarity:
Best Local Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ø
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           409
                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Unknown
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Unknown
Alignment Scores:
Pred. No.:
                                                                                                                                                                                                                   83
                                                                                                                                                                                                                                                                                                                                                                                      182
                                                                                                                                                                                   62
                                                                                                                                                                                                                                                                                                                                                                                                                                     LOCUS
DEFINITION
ACCESSION
VERSION
                                         Query Match:
DB:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          BASE COUNT
ORIGIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ORGANISM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 AUTHORS
TITLE
                                                                                                                                                                                                                                                                                                                                                                                                                      RESULT 11
AR145261
                                                                                                                                                                                                                                                                                                                                                                                                                                                                       KEYWORDS
SOURCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         REFERENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          JOURNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   FEATURES
                 Score:
                                                                                                                                                                                                                                                                                                                                                                                                     셤
                                                                                                qq
                                                                                                                 ö
                                                                                                                                 qq
                                                                                                                                                   ò
                                                                                                                                                                 qq
                                                                                                                                                                                  ò
                                                                                                                                                                                                  g
                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                 a
                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                  염
                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                    qq
                                                                                                                                                                                                                                                                                                                                                                                    οy
```

<pre>cort Similarity: 91.84% Conservative: 14 Local Similarity: 84.69% Mismatches: 13</pre>	2 Alacincintintintintintintintintintintintintinti	
Percent S. Best Local Query Mate DB: US-09-965	Db Db D D D D D D D D D D D D D D D D D	,
Pred. No.:       1.77e-63       Length:       1998         Score:       863.50       Matches:       166         Percent Similarity:       91.84%       Conservative:       14         Best Local Similarity:       84.69%       Mismatches:       13         Query Match:       64.99%       Indels:       3         DB:       6 Gaps:       1	99 5 G19Ser's Walled SCHWER CHANGE AND THE MILE AND THE M	

```
/ Acganiam ** Pearl to the 11*

/ (Ab_Xref** taxon: 11099**)

1. 385

1. 385

1. 386

1. 386

1. 386

1. 386

2. 386

2. 386

3. 386

2. 386

2. 386

2. 386

2. 386

2. 386

2. 386

2. 386

3. 386

2. 386

2. 386

2. 386

2. 386

2. 386

2. 386

2. 386

2. 386

2. 386

2. 386

2. 386

2. 386

2. 386

2. 386

2. 386

2. 386

2. 386

2. 386

2. 386

2. 386

2. 386

2. 386

2. 386

2. 386

2. 386

2. 386

2. 386

2. 386

2. 386

2. 386

2. 386

2. 386

2. 386

2. 386

2. 386

2. 386

2. 386

2. 386

2. 386

2. 386

2. 386

2. 386

2. 386

2. 386

2. 386

2. 386

2. 386

2. 386

2. 386

2. 386

2. 386

2. 386

2. 386

2. 386

2. 386

2. 386

2. 386

2. 386

2. 386

2. 386

2. 386

2. 386

2. 386

2. 386

2. 386

2. 386

2. 386

2. 386

2. 386

2. 386

2. 386

2. 386

2. 386

2. 386

2. 386

2. 386

2. 386

2. 386

2. 386

2. 386

2. 386

2. 386

2. 386

2. 386

2. 386

2. 386

2. 386

2. 386

2. 386

2. 386

2. 386

2. 386

2. 386

2. 386

2. 386

2. 386

2. 386

2. 386

2. 386

2. 386

2. 386

2. 386

2. 386

2. 386

2. 386

2. 386

2. 386

2. 386

2. 386

2. 386

2. 386

2. 386

2. 386

2. 386

2. 386

2. 386

2. 386

2. 386

2. 386

2. 386

2. 386

2. 386

2. 386

2. 386

2. 386

2. 386

2. 386

2. 386

2. 386

2. 386

2. 386

2. 386

2. 386

2. 386

2. 386

2. 386

2. 386

2. 386

2. 386

2. 386

2. 386

2. 386

2. 386

2. 386

2. 386

2. 386

2. 386

2. 386

2. 386

2. 386

2. 386

2. 386

2. 386

2. 386

2. 386

2. 386

2. 386

2. 386

2. 386

2. 386

2. 386

2. 386

2. 386

2. 386

2. 386

2. 386

2. 386

2. 386

2. 386

2. 386

2. 386

2. 386

2. 386

2. 386

2. 386

2. 386

2. 386

2. 386

2. 386

2. 386

2. 386

2. 386

2. 386

2. 386

2. 386

2. 386

2. 386

2. 386

2. 386

2. 386

2. 386

2. 386

2. 386

2. 386

2. 386

2. 386

2. 386

2. 386

2. 386

2. 386

2. 386

2. 386

2. 386

2. 386

2. 386

2. 386

2. 386

2. 386

2. 386

2. 386

2. 386

2. 386

2. 386

2. 386

2. 386

2. 386

2. 386

2. 386

2. 386

2. 386

2. 386

2. 386

2. 386

2. 386

2. 386

2. 386

2. 386

2. 386

2. 386

2. 386

2. 386
```

```
IEFCSHIPVPVRWSDMTSSHAAGREDTAVILESKAATRIDSSGERGTIJAVERKAVAFSFILL
MYSWNPLVRRICLLVLSQQPETDPSKHATYYYKGDPIGAYKDVIGRNLSELKRTGFER
LANLAKLSLSTGIWTKHTSKRIIQDCVAIGKERGNWLVNADRLISSKTGHLYTPDKGF
TLQGGYYEQLGATFTNPVMGVGTERYKLGPIVNLLLRRLKILLATAVGVSS*
12509. . 12734
DI ITDHMTEVPVITADGEVY IRNGQRGSGQPDTSAGNSMLNVLTMMYAFCESTGVPYK
          SPNRVARIHVCGDDGFLITEKGLGLKFANKGMQILHEAGKPQKITEGEKMKVAYRFED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             08-AUG-2001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                     772
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           141
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     161
                                                                                                                                                                                                                                                                                                                                                                             652
                                                                                                                                                                                                                                                                                                                                                                                                   101
                                                                                                                                                                                                                                                                                                                                                                                                                                               121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               832
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           892
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               181
                                                                                                                                                                                                                                   532
                                                                                                                                                                                                                                                                                                                                   592
                                                                                                                                                                                                                                                                    41
                                                                                                                                                                                                                                                                                                             61
                                                                                                                                                                                                                                                                                                                                                         81
                                                                                                                                                                                                                                                                               ProLysGlyProvalThrGlnMetTyrThrAsnvalAspLysAspLeuValGlyTrpGln
                                                                                                                                                                                                                                                                                                                                                                                                              122 ValThrArgHisAlaAspValIleProValArgArgArgGlyAspSerArgGlySerLeu
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       LeuSerProArgProlleSerTyrLeuLysGlySerSerGlyGlyProLeuLeuCysPro
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              GlySerValValIleValGlyArglleAsnLeuSerGlyAsp-----ThrAlaTyr
                                                                                                                                                                                                                                                                    AlaGlnGlnThrArgGlyGluGlnGlyThrGlnLysThrSerHisThrGlyArgAspLys
                                                                                                                                                                                                                                                                                                             AsnGlnValGluGlyGluValGlnIleValSerThrAlaThrGlnThrPheLeuAlaThr
                                                                                                                                                                                                                                                                                                                                                                                                                                               102 AlaProGlnGlySerArgSerLeuThrProCysThrCysGlySerSerAspLeuTyrLeu
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         833 CTTTCGCCCCGGCCCATTTCCTACTAAAGGCTCCTCTGGGGGTCCGCTGTTGTGCCCC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1 (bases 1 to 651)
Malcolm,B.A., Taremi,S.Shane., Weber,P.C. and Yao,N.
Single-chain recombinant complexes of hepatitis C virus NS3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             PAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ValAspPhelleProValGluSerLeuGluThrThrMetArgSer 196
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1 others
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             linear
                                                                                                                      Length:
Matches:
Conservative:
Mismatches:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             DNA
                                                                           g 2802 t
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            51 bp
6211338.
                                                                                                                                                                   Indels:
                                                                                                                                                                                                 US-09-965-594-22 (1-197) x AF268278 (1-12734)
                                                                                                                                                                            Gaps:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          551
Sequence 98 from patent US 62.
AR145257
AR145257.1 GI:15107124
                                                                          3293
                                                                                                                      1.83e-62
862.50
91.28%
88.72%
84.89%
                                                                             U
                                                                          2608
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Unclassified
                                                                                                                                         Percent Similarity:
Best Local Similarity:
Query Match:
DB:
                                                                             ď
                                                                           4030
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Unknown.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Unknown
                                                                                                            Scores:
                                                                                                                                                                                                                        'n
                                                                                                                                                                                                                                                                                                             42
                                                                                                                                                                                                                                                                                                                                   533
                                                                                                                                                                                                                                                                                                                                                                                                                          653
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               773
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           182
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                953
                                                                                                                                                                                                                                                                   52
                                                                                                                                                                                                                                                                                         473
                                                                                                                                                                                                                                                                                                                                                                                                   82
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     142
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     DEFINITION
ACCESSION
                                                                3'UTR
BASE COUNT
ORIGIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ORGANISM
                                                                                                                        Pred. No.:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    REFERENCE
AUTHORS
                                                                                                            Alignment
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      RESULT 14
AR145257
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             VERSION
KEYWORDS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         TITLE
                                                                                                                                  Score:
                                                                                                                                                                                                                         õ
                                                                                                                                                                                                                                             a
                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                  a
                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                           ద
                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                       a
                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                 a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               á
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                a
```

```
ซ
                        119
                                                Aliqument Scores:
                                                                                                                                                                                                                                                      82
                                                                                                                                                                                                                                                                                                                                                                                                                182
                                                                                                                                                                                                                                                                                                                                                                                                                              604
                                                                                      Query Match:
DB:
                       BASE COUNT
ORIGIN
                                                        :
⊋
                                                                                                                                                                                                                                                                                                                                                                                                                                                       Search co
 FEATURES
                                                                Score:
                                                        Pred.
                                                                                                                                            a
                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                    9
                                                                                                                                                                                                                                                                                                                               q
                                                                                                                                                                                                                                                                                                                                                             q
                                                                                                                                                                                                                                                                                                                                                                                 ŏ
                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                δ
                                                                                                                             õ
                                                                                                                                                          οŽ
                                                                                                                                                                        g
                                                                                                                                                                                         δλ
                                                                                                                                                                                                    g
                                                                                                                                                                                                                      õ
                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                    ογ
                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                              08-AUG-2001
                                                                                                                                                                                                                                                                   243
                                                                                                                                                                                                                                     AlaProGlnGlySerArgSerLeuThrProCysThrCysGlySerSerAspLeuTyrLeu 121
                                                                                                                                                                                                                                                                                                                                                                142 LeuSerProArgProlleSerTyrLeuLysGlySerSerGlyGlyProLeuLeuCysPro 161
                                                                                                                                                                                                                                                                                                                                                                                                162 AlaGlyHisAlaValGlyIlePheArgAlaAlaValSerThrArgGlyValAlaLysAla 181
                                                                                                                                                                                                        61
                                                                                                                                                                         4.1
                                                                                                                                                                         AlaGlnGlnThrArgGlyGluGlnGlyThrGlnLysThrSerHisThrGlyArgAspLys
                                                                                                                                                                                                        42 AsnGlnValGluGlyGluValGlnIleValSerThrAlaThrGlnThrPheLeuAlaThr
                                                                                                                                                                                                                1 (bases 1 to 651)
Malcolm.B.A., Taremi,S.Shane., Weber,P.C. and Yao,N.
Single-chain recombinant complexes of hepatitis C virus NS3
protease and NS4A cofactor peptide
Patent: US 6211338-A 93 03-APR-2001;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                              PAT
                                                                                                                                                                                                                                                                                                                                                                                                                                       182 ValAspPheIleProValGluSerLeuGluThrThrMetArgSer 196
                                                                      651
165
15
12
                                                                     Length:
Matches:
Conservative:
Mismatches:
Indels:
                                       u
protease and NS4A cofactor peptide
Patent: US 6211338-A 98 03-APR-2001;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Sequence 93 from patent US 6211338.
AR145252
                                       145
                                                                                                            Gaps:
                                                                                                                            US-09-965-594-22 (1-197) x AR145257 (1-651)
              Location/Qualifiers
1. .651
                             /organism="unknown"
. 188 c 199 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     GI:15107119
                                                                     7.16e-64
861.50
92.31%
84.62%
84.79%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Unclassified.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     AR145252.1
                                                                                    Percent Similarity:
Best Local Similarity:
Query Match:
DB:
                                       ď
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Unknown
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Unknown
                                                               Alignment Scores:
                                                                                                                                                                        22
                                                                                                                                                                                                                       184
                                                                                                                                                                                                                                                                     82
                                                                                                                                                                                                                                                                                   304
                                                                                                                                                                                                                                                                                                   102
                                                                                                                                                                                                                                                                                                                                  122
                                                                                                                                                                                                                                                                                                                                                 124
                                                                                                                                                                                                                                                                                                                                                                                 484
                                                                                                                                                                                                                                                                                                                                                                                                                                              604
                                                                                                                                                                                                                                                                                                                                                                                                                                                                     AR145252
LOCUS
DEFINITION
                         source
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ORGANISM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ACCESSION
VERSION
KEYWORDS
SOURCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           REFERENCE
AUTHORS
TITLE
                                                                      Pred. No.:
                                       BASE COUNT
                                                                                                                                                                                                                                                                                                                                                                                                                                                              RESULT 15
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            JOURNAL
       JOURNAL
                 FEATURES
                                                ORIGIN
                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                            셤
                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                  q
                                                                                                                                                                                                                                     Q
                                                                                                                                                                                                                                                 a
                                                                                                                                                                                                                                                                 ŏ
                                                                                                                                                                                                                                                                                  qq
                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                              qq
                                                                                                                                                                                                                                                                                                                                 Š
                                                                                                                                                                                                                                                                                                                                             a
                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                            ò
                                                                                                                                                     셤
                                                                                                                                                                        ò
                                                                                                                                                                                     a
                                                                                                                                                                                                   Š
```

```
483
                                                                                                                                                                                          183
                                                                                                                                                                                                                                                                                          243
                                                                                                                                                                                                                                                                                                                                     303
                                                                                                                                                                                                                                                                                                                                                           101
                                                                                                                                                                                                                                                                                                                                                                       121
                                                                                                                                                                                                                                                                                                                                                                                                                  122 ValThrargHisAlaAspValIleProValArgArgArgGlyAspSerArgGlySerLeu 141
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        603
                                                                                                                                                                                                                                                                      61
                                                                                                                                                                                                                          41
                                                                                                                                                                                                                                                                                                               83
                                                                                                                                                                                                                                                                                                                                                                                                                                                              102 AlaProGlnGlySerArgSerLeuThrProCysThrCysGlySerSerAspLeuTyrLeu
                                                                                                                                                                                                                                                                                                              22 AlaGlnGlnThrArgGlyGluGlnGlyThrGlnLySThrSerHisThrGlyArgAspLys
                                                                                                                                                                                                                                      42 AsnGlnValGluGlyGluValGlnIleValSerThrAlaThrGlnThrPheLeuAlaThr
                                                                                                                                                                                                                                                                               ProLysGlyProValThrGlnMetTyrThrAsnValAspLysAspLeuValGlyTrpGln
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               544 TCGGGGCACGCTGTGGGCATCTTCCGGGCTGCCGTATGCACCCGGGGGGTTGCGAAGGCG
                                                                                                                                                                               5 GlySerValValIleValGlyArglleAsnLeuSerGlyAsp-----ThrAlaTyr
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    651
165
15
12
3
                                                                                                 Conservative:
Mismatches:
                                                                             Length:
Matches:
                                   145
                                                                                                                         Indels:
                                                                                                                                      Gaps
                                                                                                                                                          US-09-965-594-22 (1-197) x AR145252 (1-651)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           completed: August 31, 2003, 00:46:41
He : 2569.57 secs
Location/Qualifiers
1..651
                       /organism="unknown"
188 c 199 g
                                                                          8.69e-64
860.50
92.31%
84.62%
84.69%
                                                                                                   Percent Similarity:
Best Local Similarity:
```

```
GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd
```

```
/ Search time 182.939 Seconds
(without alignments)
2906.924 Million cell updates/sec
                                                                                                                                          1 MKKKGSVVIVGRINLSGDTA......VAKAVDFIPVESLETIMRSP 197
nucleic search, using frame_plus_p2n model
                                  August 30, 2003, 19:13:57
                                                                                                                                                                            BLOSUM62
Xgapop 10.0, Xgapext (
Ygapop 10.0, Ygapext (
Fgapop 6.0, Fgapext (
Delop 6.0, Pgapext (
                                                                                                 US-09-965-594-22
1016
                                                                                                       Title:
Perfect score:
                                                                                                                                                                                Scoring table:
 ,
OM protein
                                                                                                                                           Sequence:
                                    Run on:
```

Total number of hits satisfying chosen parameters: length: 0 length: 2000000000 Minimum DB seq Maximum DB seq

2552756 segs, 1349719017 residues

Searched:

5105512

summaries Post-processing: Minimum Match 0% Maximum Match 100% Listing first 45 su

Command line parameters:
-MODEL-frame+\_p2n.model -DEV-xlp
-Q-/cgn2\_1/USPTO\_spool/USG09965594/runat\_29082003\_151918\_28302/app\_query.fasta\_11.2872
-Q-/cgn2\_1/USPTO\_spool/USG09965594/runat\_29082003\_151918\_28302/app\_query.fasta\_11.2872
-Q-/cgn2\_1/USPTO\_spool/USG09965594/runat\_29082003\_151918\_28302/app\_query.fasta\_11.2872
-DB-N\_Geneseq\_19JUN03 -QFMT-fasta\_P -SUFFIX-ND sum62 -TRANS-human40.cd1
-LIST-&b1x -STAFT-| FND-1 - MARTIX-DISSUM62 -TRANS-human40.cd1
-MODE-LOCAL -OUTFNT-PtO -NORM-ext -HEADS12E-500 -MINLEN-0 -MAXLEN-200000000
-GSER-USG9965594\_eCGN\_1 \_L142\_etunat\_29082003\_151918\_28302 -NCFU-6 -ICPG-3
-NO\_MMAP -LANGEQUERY -NEG\_SCORES- -WAIT -DSPBLOCK=100 -LONGLOG
-EDV\_TIMBOOT-120 -WARN\_TIMEOUT-30 -THREDS-1 -XGAPOP-10 -XGAPEXT-0.5 -FGAPOP-6

Database :

1: /SIDS1/gcgdata/geneseq/geneseqn-emb1/Na1980.DAT:\*
2: /SIDS1/gcgdata/geneseq/geneseqn-emb1/Na1981.DAT:\*
4: /SIDS1/gcgdata/geneseq-emb1/Na1981.DAT:\*
4: /SIDS1/gcgdata/geneseq-emb1/Na1982.DAT:\*
5: SIDS1/gcgdata/geneseq-emb1/Na1984.DAT:\*
5: SIDS1/gcgdata/geneseq-emb1/Na1985.DAT:\*
7: /SIDS1/gcgdata/geneseq-emb1/Na1985.DAT:\*
7: /SIDS1/gcgdata/geneseq-emb1/Na1985.DAT:\*
8: /SIDS1/gcgdata/geneseq-emb1/Na1986.DAT:\*
8: /SIDS1/gcgdata/geneseq-emb1/Na1989.DAT:\*
10: /SIDS1/gcgdata/geneseq-emb1/Na1999.DAT:\*
11: /SIDS1/gcgdata/geneseq-geneseq-emb1/Na1990.DAT:\*
12: /SIDS1/gcgdata/geneseq-geneseq-emb1/Na1991.DAT:\*
13: /SIDS1/gcgdata/geneseq-geneseq-emb1/Na1992.DAT:\*
14: /SIDS1/gcgdata/geneseq-geneseq-emb1/Na1993.DAT:\*
15: /SIDS1/gcgdata/geneseq-geneseq-emb1/Na1995.DAT:\*
16: /SIDS1/gcgdata/geneseq-geneseq-emb1/Na1995.DAT:\*
17: /SIDS1/gcgdata/geneseq-geneseq-emb1/Na1995.DAT:\*
18: /SIDS1/gcgdata/geneseq-geneseq-emb1/Na1995.DAT:\*
19: /SIDS1/gcgdata/geneseq-geneseq-emb1/Na1995.DAT:\*
19: /SIDS1/gcgdata/geneseq-geneseq-emb1/Na1999.DAT:\*
19: /SIDS1/gcgdata/geneseq-geneseq-emb1/Na1999.DAT:\*
20: /SIDS1/gcgdata/geneseq-geneseq-emb1/Na1999.DAT:\*
21: /SIDS1/gcgdata/geneseq-geneseq-emb1/Na1999.DAT:\*
22: /SIDS1/gcgdata/geneseq-geneseq-emb1/Na1999.DAT:\*
23: /SIDS1/gcgdata/geneseq-geneseq-emb1/Na1999.DAT:\*
24: /SIDS1/gcgdata/geneseq-geneseq-emb1/Na1999.DAT:\*
25: /SIDS1/gcgdata/geneseq-geneseq-emb1/Na1999.DAT:\*
26: /SIDS1/gcgdata/geneseq-geneseq-emb1/Na1999.DAT:\*
27: /SIDS1/gcgdata/geneseq-geneseq-emb1/Na1999.DAT:\*
28: /SIDS1/gcgdata/geneseq-geneseq-geneseq-emb1/Na1999.DAT:\*
28: /SIDS1/gcgdata/geneseq-geneseq-geneseq-emb1/Na1999.DAT:\*
27: /SIDS1/gcgdata/geneseq-genes /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:\*/SIDS1/gcgdata/geneseq-emb1/NA2001B.DAT:\*/SIDS1/gcgdata/geneseqn-emb1/NA2001B.DAT:\*/SIDS1/gcgdata/geneseqn-emb1/NA2002.DAT:\*/SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2003.DAT:\* N\_Geneseq\_19Jun03:\* ..

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed,

and is derived by analysis of the total score distribution.

SUMMARIES

		•				COLUMNIA	
Resul	1¢	ő	lery				
N :	Scor	a :	Match	Length	DB	£	Description
	10	-	٠.	594		AAA73334	atitis
	10		4.60	594		AAA73333	Titis C
	٥		6.7	594		AAA73332	itis C
	o			594		AAA73331	tis C v
	σ,			594		AAA73330	itis C
	on c			594		AAA73335	itis C v
	non			9 60		AAA73328	Hepatitis C virus
	882		6.9	12734		ABA95615	neric BVDV/
	0 875			1998		AAX80355	84A-NS3
	1 872			1998		AAX80359	NS4A-NS3
	3 86.1			1448 653		AAX80334 bby80345	
	868			1998		AAX80343	NS4A-NS3
	5 867			1998		AAX80353	_
. • •	865			651		AAX80349	NS4A-NS3
•	7 864		2.5	612		ABX15706	i-viral syn
	808		7.5	1001		AAXBU344 aax80357	
	9004		10.5	1998		AAX80337	
	1 863			2013		AAX80360	NS4A-NS3
. • •	2 861		4.8	651		AAX80348	HCV NS4A-NS3 COMPI
	860		7.	169		AAX80343	NS4A-NS3
	860		. 7	2016		AAX80361	
. , ,	8		94.6	648		AAX80365	NS4A-NS3
. •	2		4.4	648		AAX80363	NS4A-NS3
	857		4. 4	650		AAX80347	
	0 856		. 7	651		AAX80331	
	1		34.1	648		AAXB0362	O
	2 853			650		AAX80346	NS4A-NS3
• •	ממא מ		. c	129		AAXBU350	υç
	, w			1933		AAX23258	NS3 DNA
. •	6 848			9646		AAV59361	atitis C
	7 848			9646		ABK87285	odin
• •	8 848		33.5	12980		AAV59364	epatitis C
	948		•	16622		ABNO / 200 AB236212	Nucleotide segment
	1 844			5300		AAN92097	3 8
-	2 844			5360		AAN90327	s c vir
•	3 844			6905		AAN92103	ed open rea
- •	45 844	กูเก	83.1	7310	25	AANYZIO6 AANYO336	Composite benatiti
						SHAGMAN TA	
RESU						CINTERIOR	
AAA73	3334						
ឧះ	AAA73334	S.	tandard	3; DNA;	594	BP,	
<b>4</b> 2 3	AAA73334	::					
<b>4</b> 5	19-DEC-2	2000	(f11	first entry	( <del>/</del>		
E X	Hepatiti	ပ	virus	ž	m	fusion protease coding	sequence #7.
×			,				;
K W	liver fat	failure;		protease; liver can	> ຍ	<pre>irai replication; chrobic if; mutant; mutein; ds.</pre>	ic liver disease;
XX							

Location/Qualifiers Hepatitis C virus. Synthetic. Key XX SS XX

301 CAGGCTCCGCAGGGTTCCCGTTCCCTGACCCCGTGCACCTGCGGTTCCTCCGACCTGTAC 360

a

```
The present sequence is the coding sequence for a mutated version of a fusion protein created using the Hepatitis C virus (HCV) NS3 and NS4A protease enzymes. These proteins are both sescential for the replication of the virus, acting to cleave its replicative proteins from the polyprotein produced from the HCV genome. Inhibitors of the two proteins should be effective as antiviral treatments of HCV infection. This is useful as HCV can lead to chronic liver disease such as cirrhosis, liver failure and Niver cancer. The present invention concerns a number of NS3 mutants and NS3-NS4A fusion proteins which can be used to identify inhibitors of this type, as well as enabling structural studies of the protease and protease inhibitor complexes. The protein produced from this sequence contains the alpha-helix0-7 variant.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 240
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              101 GlnAlaProGlnGlySerArgSerLeuThrProCysThrCysGlySerSerAspLeuTyr 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SerProLysGlyProValThrGlnMetTyrThrAsnValAspLysAspLeuValGlyTrp 100
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           20
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ThrserlleasnGlyvalLeuTrpThrValTyrHisGlyAlaGlyThrArgThr1leAla 80
                                                                                                                                                                                                                                                                          Modified hepatitis C virus (HCV) NS3 protease comprising at least 1 substitution of a hydrophobic alpha-helix 0 amino acid to a hydrophilic amino acid, useful for screening inhibitors that may treat hepatitis C
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         MetLysLysLysGlySerValValIleValGlyArgIleAsnLeuSerGlyAspThrAla
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           LysAsnGlnValGluGlyGluValGlnIleValSerThrAlaThrGlnThrPheLeuAla
                           /product- "NS4A-NS3 fusion protein #7"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Sequence 594 BP; 105 A; 192 C; 151 G; 146 T; 0 other;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            594
0
0
0
0
                                                                                                                                                                                                       Goldfarb V;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Length:
Matches:
Conservative:
Mismatches:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Indels:
Gaps:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                US-09-965-594-22 (1-197) x AAA73334 (1-594)
                                                                                                                                                                                                       Shang Y,
                                                                                                                                                                          8
                                                                                                                                                                                                                                                                                                                                                    Claim 26; Fig 17; 66pp; English
                                                                                                                                                                        (BRIM ) BRISTOL-MYERS SQUIBB
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            3.95e-86
1016.00
100.00%
100.00%
                                                                                                                                                                                                       Weinheimer S,
                                                                                                                                             990S-0115271.
                                                                                                                  06-JAN-2000; 2000WO-US00345
                                                                                                                                                                                                                                  WPI; 2000-465976/40.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Percent Similarity:
Best Local Similarity:
                                                                                                                                                                                                                                                  P-PSDB; AAB15225
                                                        WO200040707-A1
                                                                                                                                                                                                        Wittekind M,
                                                                                                                                              08-JAN-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Alignment Scores:
                                                                                     13-JUL-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      21
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              4,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Query Match:
DB:
 CDS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              õ
```

```
The present sequence is the coding sequence for a mutated version of a fusion protein created using the Hepatitis C virus (HCV) NS3 and NS4A protease enzymes. These proteins are both essential for the replication of the virus, acting to cleave its replicative proteins from the polyprotein produced from the HCV genome. Inhibitors of the two proteins should be effective as antiviral treatments of HCV infection. This is useful as HCV can lead to chronic liver disease such as cirrhosis, liver failure and liver cancer. The present invention concerns a number of NS3 mutants and NS3-NS4A fusion proteins which can be used to identify inhibitors of this type, as well as enabling structural studies of the protease and protease:inhibitor complexes. The protein produced from this sequence contains the alpha-helix0-7 variant.
                                                                                       480
                                                                                                                     180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Modified hepatitis C virus (HCV) NS3 protease comprising at least 1 substitution of a hydrophobic alpha-helix 0 amino acid to a hydrophilic amino acid, useful for screening inhibitors that may treat hepatitis C
                                                        161 ProalaGlyHisalavalGlyIlePheArgAlaalaValSerThrArgGlyValAlaLys
                                                                                                                                                    481 CCGGCTGGTCACGCTGTTGGTATCTTCCGTGCTGTTTCCACCCGTGGTGTTGCTAAA
                LeuValThrArgHisAlaAspVallleProValArgArgArgGlyAspSerArgGlySer
                                                                                                                                                                                                                                                                                                                                                                                              Hepatitis; NS3 protease; viral replication; chronic liver disease; liver failure; liver cancer; mutant; mutein; ds.
                                                                                                                                                                               Hepatitis C virus NS4A-NS3 fusion protease coding sequence #6.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         /*tag* a
/product= "NS4A-NS3 fusion protein #6"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Goldfarb V;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Zhang Y,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Claim 26; Fig 16; 66pp; English.
                                                                                                                                                                                                                                                                             BP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          (BRIM ) BRISTOL-MYERS SQUIBB
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Weinheimer S,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                06-JAN-2000; 2000WO-US00345
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            99US-0115271
                                                                                                                                                                                                                                                                             AAA73333 standard; DNA; 594
                                                                                                                                                                                                                                                                                                                                         (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    WPI; 2000-465976/40.
P-PSDB; AAB15224.
                                                                                                                                                                                                                                                                                                                                                                                                                                               Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    WO200040707-A1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Wittekind M,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             08-JAN-1999;
                                                                                                                                                                                                                                                                                                                                       19-DEC-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 13-JUL-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                               Synthetic.
                              361
                                                                                                                                                                                                                                                                                                       AAA73333;
121
                                                                                                                                                                                                                                                                               δ
                                                                                                                                                    임
ŏ
                         q
                                                           ò
                                                                                        Q
```

us-09-965-594-22.rng

```
301 CAGGCTCCGCAGGGTTCCCGTTCCCTGACCCCGTGCACCTGCGGTTCCTCCGACCTGTAC 360
                                                                                                                                                                                                                                                                                                                                                                                                              LeuValThrArgHisAlaAspValIleProValArgArgArgGlyAspSerArgGlySer 140
                                                                                                                                                                                                                                 Percent Similarity:
Best Local Similarity:
                                                                             P-PSDB; AAB15223
     40200040707-A1
                                      38-JAN-1999;
                                                                                                                                                                                                                 Alignment Scores:
                13-JUL-2000
                                                                                                                                                                                                                                                                        Н
                                                                                                                                                                                                                                                                                                                              121
                                                                                                                                                                                                                                                                                                                                                     181
                                                                                                                                                                                                                                                                                                                    41
                                                                                                                                                                                                                                                                                                                                                                                                              121
                                                                                                                                                                                                                                                                                                                                          61
                                                                                                                                                                                                                                                                                                                                                                 81
                                                                                                                                                                                                                                             Query Match:
                                                                                                                                                                                                                      Pred. No.:
                                                                                                                                                                                                                           Score:
                                                                                                                                                                                                                                                                                                                                                   a
                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                 q
ò
                                                                                                                                                                                                                                                                                  a
                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                       a
                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                       ŏ
                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                   240
                                                                                                                                                              100
                                                                                                                                                                         300
                                                                                                                                                                                               360
                                                                                                                                                                                                           140
                                                                                                                                                                                                                      420
                                                                                                                                                                                                                                 160
                                                                                                                                                                                                                                            480
                                                                                                                                                                                                                                                       180
                                                                                                                                                                                     GlnalaProGlnGlySerArgSerLeuThrProCysThrCysGlySerSerAspLeuTyr 120
                                                                                                       120
                                                                                                                  9
                                                                                                                                         80
                                                                       20
                                                                                 9
                                                                                             40
                                                                                            SerProLysGlyProValThrGlnMetTyrThrAsnValAspLysAspLeuValGlyTrp
                                                                                                                                                                                           LeuValThrArgHisAlaAspValIleProValArgArgArgGlyAspSerArgGlySer
                                                                                                                                                                                                                MetLysLysLysGlySerValValIleValGlyArgIleAsnLeuSerGlyAspThrAla
                                                                             tis; NS3 protease; viral replication; chronic liver disease; failure; liver cancer; mutant; mutein; ds.
                                                                                                                                                                                                                                                                            Hepatitis C virus NS4A-NS3 fusion protease coding sequence #5.
                                                                                                                                                                                                                                                                                                                                                                                                       /*tag= a
/product= "NS4A-NS3 fusion protein #5"
     other;
                     594
196
0
0
0
                    Length:
Matches:
Conservative:
Mismatches:
    0
    Ë
                                           Indels:
Gaps:
    147
                                                           (1-594)
    3
                                                                                                                                                                                                                                                                                                                                                                                             Location/Qualifiers
    152
    <u>ن</u>
                                                           US-09-965-594-22 (1-197) x AAA73333
                                                                                                                                                                                                                                                                                                              BP
                    1.44e-85
1010.00
99.49%
99.49%
21
    104 A; 191
                                                                                                                                                                                                                                                                                                              DNA; 594
                                                                                                                                                                                                                                                                                                                                    19-DEC-2000 (first entry)
                                                                                                                                                                                                                                                                                                              standard;
                                                                                                                                                                                                                                                                                                                                                                            Hepatitis C virus
                                      Best Local Similarity:
    BP;
                                                                                                                                                                                                                                                                                                                                                          Hepatitis; NS3
                                Percent Similarity:
    594
                Alignment Scores
                                                                                                                                                                                                                                                                                                                                                                                 Synthetic
                                                                                             21
                                                                                                                                                                                                                                                                 481
     Sequence
                                                                                                                                                    181
                                                                                                                                                                         241
                                                                                                                                                                                                                     361
                                                                                                                                                                                                                                 141
                                                                                                                                                                                                                                           421
                                                                                                                                                                                                                                                       161
                                                                                                                                                                                                                                                                                                              AAA73332
                                                                                                                                                                                    101
                                                                                                                                                                                                301
                                                                                                                                                                                                           121
                                                                                                                                                                                                                                                                            181
                                                                                                                                         61
                                                                                                                                                              81
                                                                                                                                                                                                                                                                                                                        AAA73332
                                          Query Match:
DB:
                      :.
₽
                                                                                                                                                                                                                                                                                                  ö
                                                                                g
                                                                                            ò
                                                                                                                                                  g
                                                                                                                                                              ò
                                                                                                                                                                       g
                                                                                                                                                                                  οχ
                                                                                                                                                                                              염
                                                                                                                                                                                                          ò
                                                                                                                                                                                                                  셤
                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                          셤
                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                               a
                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                       qq
                                                                                                     g
                                                                                                                ò
                                                                                                                            g
                                                                                                                                         ò
```

```
The present sequence is the coding sequence for a mutated version of a fusion protein created using the Hepatitis C virus (HCV) NS3 and NS4A protease enzymes. These proteins are both essential for the replication of the virus, acting to cleave its replicative proteins from the polyprotein produced from the HCV genome. Inhibitors of the two proteins should be effective as antiviral treatments of HCV infection. This is useful as HCV can lead to chronic liver disease such as cirrhosis, liver failure and Niver cancer. The present invention concerns a number of NS3 mutants and NS3-NS4A fusion proteins which can be used to identify inhibitors of this type, as well as enabling structural studies of the protease:inhibitor complexes. The protein produced from this sequence contains the alpha-helix0-1 variant.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SerProLysGlyProValThrGlnMetTyrThrAsnValAspLysAspLeuValGlyTrp 100
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        GlnAlaProGlnGlySerArgSerLeuThrProCysThrCysGlySerSerAspLeuTyr 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   LysAsnGlnValGluGlyGluValGlnIleValSerThrAlaThrGlnThrPheLeuAla 60
                                                                                                                                                                                                                                                                                                                              Modified hepatitis C virus (HCV) NS3 protease comprising at least 1 substitution of a hydrophobic alpha-helix 0 amino acid to a hydrophilic amino acid, useful for screening inhibitors that may treat hepatitis C
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    MetLysLysLysGlySerValVallleValGlyArglleAsnLeuSerGlyAspThrAla
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             TyralaClnGlnThrArgGlyGluGlnGlyThrGlnLysThrSerHisThrGlyArgAsp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ThrSerIleAsnGlyValLeuTrpThrValTyrHisGlyAlaGlyThrArgThrIleAla
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1 ATGAAAAAAAAGGATCCGTTGTTATCGTCGGCCGTATCAACCTGTCCGGTGACACCGCT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  AAAAACCAGGTTGAAGGTGAAGTTCAGATGGTTTCCACCGCTACCCAGACCTTCCTGGCT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ACCTCCATCAACGGTGTTCTGTGGACCGTTTACCACGGTGCTGGTACCCGTACCATCGCT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           594
193
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Sequence 594 BP; 105 A; 189 C; 153 G; 147 T; 0 other;
                                                                                                                                                                                  Goldfarb V;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Length:
Matches:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              US-09-965-594-22 (1-197) x AAA73332 (1-594)
                                                                                                                                                                               Zhang Y,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Claim 26; Fig 15; 66pp; English.
                                                                                                                         S
                                                                                                                     (BRIM ) BRISTOL-MYERS SQUIBB
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        3.64e-84
995.00
98.98%
97.97%
                                                                                                                                                                                  Wittekind M, Weinheimer S,
06-JAN-2000; 2000WO-US00345
                                                         99US-0115271
                                                                                                                                                                                                                                     WPI; 2000-465976/40.
```

180

```
Alignment Scores:
                             Pred. No.:
                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                      9
                                                                                                                                                                                                                                                                                                                                                                                                                                                           ó
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             οy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           fusion protein created using the Hepatitis C virus (HCV) NS3 and NS4A protease enzymes. These proteins are both essential for the replication of the virus, acting to cleave its replicative proteins from the polyprotein produced from the HCV genome. Inhibitors of the two proteins should be effective as antiviral treatments of HCV infection. This is useful as HCV can lead to chronic liver disease such as cirrhosis, liver failure and liver cancer. The present invention concerns a number of NS3 mutants and NS3-NS4A fusion proteins which can be used to identify inhibitors of this type, as well as enabling structural studies of the protease and protease:inhibitor complexes. The protein produced from this sequence contains the alpha-helix0-1 variant.
                                                                                                                                                                                                                             540
                                                                            160
480
                                                                                                                                                                           180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Modified hepatitis C virus (HCV) NS3 protease comprising at least 1 substitution of a hydrophobic alpha-helix 0 amino acid to a hydrophilic amino acid, useful for screening inhibitors that may treat hepatitis C
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    present sequence is the coding sequence for a mutated version of a
                                                                         ProAlaGlyHisAlaValGlyIlePheArgAlaAlaValSerThrArgGlyValAlaLys
                                                                                                                                                                                                                           Hepatitis; NS3 protease; viral replication; chronic liver disease;
liver failure; liver cancer; mutant; mutein; ds.
                                                                                                                                                                                                                                                                               AlaValAspPheIleProValGluSerLeuGluThrThrMetArgSerPro 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Hepatitis C virus NS4A-NS3 fusion protease coding sequence #4.
                                                                                                                                                                                                                                                                                                    /*tag= a
/product= "NS4A-NS3 fusion protein #4"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Goldfarb V;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ۲,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Location/Qualifiers
1..594
/*tag= a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    2hang
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Claim 26; Fig 14; 66pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  (BRIM ) BRISTOL-MYERS SQUIBB CO
                                                                                                                                                                                                                                                                                                                                                                                                                                   BP
                                                                                                                                                                                                                                                                                                                                                                                                                                   AAA73331 standard; DNA; 594
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   06-JAN-2000; 2000WO-US00345
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Wittekind M, Weinheimer S,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 WPI; 2000-465976/40.
P-PSDB; AAB15222.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 W0200040707-A1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  08-JAN-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  19-DEC-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 13-JUL-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Synthetic
                                                                                                                                                                        161
                                                                                                                                                                                                                                                                             181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   AAA73331;
                       361
                                                                                                                                                                                                                                                                                                                                                                         RESULT 4
AAA73331
AAA73331
AAA73331
AAA73331
AAA73331
AAA73331
AAA733331
AAAA733331
AAA733331
AAA733331
AAA733331
AAA733331
AAA733331
AAA733331
AAA733331
AAA733331
AAA733331
AAAA733331
AAA733331
AAA73331
AAA733331
AA
                                                                                                                                                                                                                                                                                                               8
                                                                         ò
                                                                                                              g
                                                                                                                                                           ò
                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                             ò
```

594 BP; 105 A; 187 C; 155 G; 147 T; 0 other;

Sequence

```
160
                                                                                                                                                                                                                                                                                                                                                                                                            ProAlaGlyHisAlaValGlyIlePheArgAlaAlaValSerThrArgGlyValAlaLys 180
                                                                                                                                                                                                                                240
                                                                                                                                                                                                                                                   100
                                                                                                                                                                                                                                                                                          120
                                                                                                                                                                                                                                                                                                            360
                                                                                                                                                                                                                                                                                                                                140
                                                                                                                                                                                                                                                                                                                                                   420
                                                                                                                                                                                                                                                                                                                                                                                         480
                                                                                                                                                                                                                                                                                                                                                                                                                                481 CCGCCTCGTCACGCTGTTGGTATCTTCCGTGCTGTTTGCACCCGTGGTGTTGCTAAA 540
                                                                                                                                         180
                                                                                                                                                                                                                                                                     300
                                                                                                                                                                       9
                                                                                                                                                                                                            80
                                                                                                                               0.4
                                                                                                           9
                                                                                       GlaAlaProGlaGlySerArgSerLeuThrProCysThrCysGlySerSerAspLeuTyr
                                                                                                                                                                                                                                                                                                                                            301 CAGGCTCCGCAGGGTTCCCGTTCCCTGACCCCGTGCACCTGCGGTTCCTCCGACCTGTAC
                                                                                                                                                                                                                                                                                                                                                                        LeuLeuSerProArgProlleSerTyrLeuLysGlySerSerGlyGlyProLeuLys
                                                                                                                               21 TyralaGlnGlnThrArgGlyGluGlnGlyThrGlnLysThrSerHisThrGlyArgAsp
                                                                                                                                                                      LysAsnGlnValGluGlyGluValGlnIleValSerThrAlaThrGlnThrPheLeuAla
                                                                                                                                                                                                                                                                                                                                LeuValThrArgHisAlaAspValIleProValArgArgArgGlyAspSerArgGlySer
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Hepatitis; NS3 protease; viral replication; chronic liver disease;
liver failure; liver cancer; mutant; mutein; ds.
                                                                                                                                                                                                                                                                                                                                                                                                                                                   Hepatitis C virus NS4A-NS3 fusion protease coding sequence #3.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           /*tag= a
/product= "NS4A-NS3 fusion protein #3"
 594
190
2
0
0
        Matches:
Conservative:
Mismatches:
 Length:
                                         Indels:
                                                  Gaps:
                                                                    US-09-965-594-22 (1-197) x AAA73331 (1-594)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Location/Qualifiers
1..594
/*tag= a
9.23e-83
980.00
97.46%
96.45%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             AAA73330 standard; DNA; 594
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Hepatitis C virus.
Synthetic.
                 Percent Similarity:
Best Local Similarity:
Query Match:
DB:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        WO200040707-A1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     19-DEC-2000
                                                                                                                                                                      41
                                                                                                                                                                                         121
                                                                                                                                                                                                                                                                                                                                                                       141
                                                                                                                                                                                                                                                                                                                                                                                                             161
                                                                                                                                                                                                                                                                                        101
                                                                                                                                                                                                                                                                                                                                121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 AAA73330;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          RESULT 5
```

```
08-JAN-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Alignment Scores:
                                                                                                                                                                                                        19-DEC-2000
                                                                                                                                                                                                                                                                                                                                                                                                                             13-JUL-2000.
                                                                                                                                                                                                                                                          Hepatitis;
                                                                                                                                                                                                                                                                                                            Synthetic
                     161
                                                                                                                                                                              AAA73335;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Pred. No.:
Score:
                                                                                                                              RESULT
                                                                                                                                                      a
                     ŏ
                                              qq
                                                                        ò
                                                                                                                                                                                                                           The present sequence is the coding sequence for a mutated version of a fusion protein created using the Hepatitis C virus (HCV) NS3 and NS4A protease enzymes. These proteins are both essential for the replication of the virus, acting to cleave its replicative proteins from the Polyprotein produced from the HCV genome. Inhibitors of the two proteins should be effective as antiviral treatments of HCV infection. This is useful as HCV can lead to chronic liver disease such as cirrhosis, liver failure and NS3-NS4A fusion proteins which can be used to identify inhibitors of this type, as well as enabling structural studies of the protease and protease inhibitor complexes. The protein produced from this sequence contains the alpha-helix0-l variant.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            SerProLysGlyProValThrGlnMetTyrThrAsnValAspLysAspLeuValGlyTrp 100
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 300
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CCGGCTCCGCAGGCTTCCCGTTCCCTGACCCCGTGCACCTCCGCGGTTCCTCCGCGACCTGAAC 360
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      TACGCTCAGCAGACTCGAGGTGAGGAGGGTTGCCAAGAAACCTCCCAGACGGGTCGTGAC 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               240
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           101 GlnAlaProGinGlySerArgSerLeuThrProCysThrCysGlySerSerAspLeuTyr 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ThrSerIleAsnGlyValLeuTrpThrValTyrHisGlyAlaGlyThrArgThrIleAla 80
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             TyralacinchnthrargciyclucinciythrcinnysthrSerHisThrciyArgAsp 40
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                LysAsnGlnValGluGlyGluValGlnIleValSerThrAlaThrGlnThrPheLeuAla 50
                                                                                                                                         Modified hepatitis C virus (HCV) NS3 protease comprising at least 1 substitution of a hydrophobic alpha-helix 0 amino acid to a hydrophilic amino acid, useful for screening inhibitors that may treat hepatitis C
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 MetLysLysLysLysGlySerValValIleValGlyArgIleAsnLeuSerGlyAspThrAla
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            BP; 103 A; 186 C; 156 G; 149 T; 0 other;
                                                                                                                                                                                                                                                                                                                                                                                                                                          594
187
2
8
0
0
                                                                           Goldfarb V;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Gaps:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (1-594)
                                                                           Zhang Y,
                                                                                                                                                                                                        26; Fig 13; 66pp; English.
                                                  ဗ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      US-09-965-594-22 (1-197) x AAA73330
                                                  (BRIM ) BRISTOL-MYERS SQUIBB
                                                                                                                                                                                                                                                                                                                                                                                                                                          3.59e-81
963.00
95.94%
94.92%
                                                                          Weinheimer S,
06-JAN-2000; 2000WO-US00345
                                                                                                     2000-465976/40.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Percent Similarity:
Best Local Similarity:
                                                                                                                   P-PSDB; AAB15221
                                                                           Wittekind M,
                                                                                                                                                                                                                                                                                                                                                                                                      Sequence 594
                                                                                                                                                                                                                                                                                                                                                                                                                                Aliqument Scores:
                          08-JAN-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                21
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  241
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       81
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      301
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Query Match:
DB:
                                                                                                                                                                                                          Claim
                                                                                                                                                                                                                                                                                                                                                                                                                                              ..
Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ď
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
```

g

Š

à g Š ò

Ω

à

```
The present sequence is the coding sequence for a mutated version of a fusion protein created using the Hepatitis C virus (HCV) NS3 and NS4A protease enzymes. These proteins are both essential for the replication of the virus, acting to cleave its replicative proteins from the polyprotein produced from the HCV genome. Inhibitors of the two proteins should be effective as antiviral treatments of HCV infection. This is useful as HCV can lead to chronic liver disease such as cirrhosis, liver diallure and liver cancer. The present invention concerns a number of NS3 mutants and NS3-NS4A fusion proteins which can be used to identify inhibitors of this type, as well as enabling structural studies of the protease and protease:inhibitor complexes. The protein produced from this
                                                                                                                             180
                                                                                                                                                                                       540
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Modified hepatitis C virus (HCV) NS3 protease comprising at least 1 substitution of a hydrophobic alpha-helix 0 amino acid to a hydrophilic amino acid, useful for screening inhibitors that may treat hepatitis {\sf C}
ProalaglyHisAlaValGlyIlePheArgAlaAlaValSerThrArgGlyValAlaLys
                                                                                                                                                          NS3 protease; viral replication; chronic liver disease; ure; liver cancer; mutant; mutein; ds.
                                                                                                                                                                                                                                                    181 AlaValAspPheIleProValGluSerLeuGluThrThrMetArgSerPro 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Hepatitis C virus NS4A-NS3 fusion protease coding sequence #8.
                                                                                                                                                                                                                                                                                       sequence contains the alpha-helix0 wild-type sequence.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          /product- "NS4A-NS3 fusion protein #8"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 594
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Sequence 594 BP; 98 A; 189 C; 153 G; 154 T; 0 other;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Goldfarb
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Length:
Matches:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 2hang Y,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Location/Qualifiers
1..594
/*tag= a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Disclosure; Fig 18; 66pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        8
                                                                                                                                                                                                                                                                                                                                                                                                                                        ВР
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (BRIM ) BRISTOL-MYERS SQUIBB
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ŝ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          06-JAN-2000; 2000WO-US00345
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           99US-0115271
                                                                                                                                                                                                                                                                                                                                                                                                                                        AAA73335 standard; DNA; 594
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Wittekind M, Weinheimer
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           2000-465976/40.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           P-PSDB; AAB15226.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  liver failure;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       WO200040707-A1
```

```
81
                                                                                                                                                                                                                                                                             Query Match:
                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                 QQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           셤
 ò
                                                                                                                                                                                                                                                                                                                                                           q
                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                       \delta
                                                                                                                                                                                                                                                                                                                                                                                                                     a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ó
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             à
                                                                                                                                                                                                                                                                                                                  à
                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                 140
                                                                                                                                                                                                                                                              160
                                                                                                                                                         240
                                                                                                                                                                        100
                                                                                                                                                                                     300
                                                                                                                                                                                                   101 GlnAlaProGlnGlySerArgSerLeuThrProCysThrCysGlySerSerAspLeuTyr 120
                                                                                                                                                                                                                  360
                                                                                                                                                                                                                                               420
                                                                                                                                                                                                                                                                             480
                                                                                                                                                                                                                                                                                            180
                                                                                                                                                                                                                                                                                                          540
                                                                                                                            180
                                                   50
                                                                  9
                                                                                 40
                                                                                                              9
                                                                                                                                           80
                                                                                                                                                                      LeuLeuSerProArgProIleSerTyrLeuLysGlySerSerGlyGlyProLeuLeuCys
                                                                                                                                                                                                                                                                                                   LeuValThrArgHisAlaAspValIleProValArgArgArgGlyAspSerArgGlySer
                                                                                                                                                                                                                                                                                           ProAlaGlyHisAlaValGlyIlePheArgAlaAlaValSerThrArgGlyValAlaLys
                                                                                         ThrSerIleAsnGlyValLeuTrpThrValTyrHisGlyAlaGlyThrArgThrIleAla
                                                                                                                                                         181 ACCTGCATCAACGGTGTTTGCTGGACCGTTTACCACGGTGCTGGTACCCGTACCATCGCT
                                                                                                                                                                                                                                                                             CTGCTGTCCCCGCGTCCGATCTCCTACCTGAAAGGTTCCTCCGGTGGTCCGCTGCTGTGC
                                                                                 TyralaGlnGlnThrargGlyGluGlnGlyThrGlnLysThrSerHisThrGlyArgAsp
                                                                                                              LysAsnGlnValGluGlyGluValGlnIleValSerThrAlaThrGlnThrPheLeuAla
                                                                                                                     MetLysLysLysClySerValValIleValGlyArgIleAsnLeuSerGlyAspThrAla
                                                           1 ATGAAAAAAAAAGGATCCGTTGTTATCGTCGCCGTATCAACCTGTCCGGTGACACTGT
                                                                                                                                                                                                                                                                                                                                                                                                                             ltis; NS3 protease; viral replication; chronic liver disease;
failure; liver cancer; mutant; mutein; ds.
                                                                                                                                                                                                                                                                                                                         AlaValAspPheIleProValGluSerI.euGluThrThrMetArgSerPro 197
                                                                                                                                                                                                                                                                                                                                Hepatitis C virus NS4A-NS3 fusion protease coding sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               "NS4A-NS3 fusion protein #2"
0050
 Conservative
       Mismatches:
Indels:
                                     US-09-965-594-22 (1-197) x AAA73335 (1-594)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Location/Qualifiers
1..588
                                                                                                                                                                                                                                                                                                                                                                   ВР
                                                                                                                                                                                                                                                                                                                                                                   AAA73329 standard; DNA; 588
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           06-JAN-2000; 2000WO-US00345
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        /product *
                                                                                                                                                                                                                                                                                                                                                                                                 (first entry)
93.91%
93.91%
92.13%
                                                                                                                                                                                                                                                                                                                                                                                                                                                  Hepatitis C virus. Synthetic.
       Best Local Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                              Hepatitis; NS3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              WO200040707-A1
Percent Similarity
                                                                                                                                                                                                                                                                                                                                                                                               19-DEC-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             13-JUL-2000
                                                                                                                                                                                                                                                                           421
                                                                                                              41
                                                                                                                                           61
                                                                                                                                                                        81
                                                                                                                                                                                     241
                                                                                                                                                                                                                  301
                                                                                                                                                                                                                                                              141
                                                                                                                                                                                                                                                                                           161
                                                                                                                                                                                                                                                                                                         481
                                                                                                                                                                                                                                                                                                                        181
                                                                                                                                                                                                                                                                                                                                       541
                                                                                                                                                                                                                                                                                                                                                                                  AAA73329
                 Query Match:
                                                                                                                                                                                                                                                                                                                                                                                                                                     liver
                                                                                                                                                                                                                                                                                                                                                   Key
                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                       q
                                                                                                                                                      a
                                                                                                                                                                                  Q
                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                           à
                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                     g
                                                   ð
                                                                g
                                                                                ò
                                                                                             g
                                                                                                            ò
                                                                                                                          g
                                                                                                                                          ò
                                                                                                                                                                      ŏ
                                                                                                                                                                                                  οŽ
```

```
fusion protein created using the Hepatitis C virus (HCV) NS3 and NS4A protease enzymes. These proteins are both essential for the replication of the virus, acting to cleave its replicative proteins from the polyprotein produced from the HCV genome. Inhibitors of the two proteins should be effective as antiviral treatments of HCV infection. This is useful as HCV can lead to chronic liver disease such as cirrhosis, liver failure and liver cancer. The present invention concerns a number of NS3 mutants and NS3-NS4A fusion proteins which can be used to identify inhibitors of this type, as well as enabling structural studies of the protease and protease:inhibitor complexes. The protein produced from this sequence contains the alpha-helix0-1 variant.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               101 GlnalaProGlnGlySerArgSerLeuThrProCysThrCysGlySerSerAspLeuTyr 120 [1111] [1111] [1111] [1111] [1111] [1111] [1111] [1111] [1111] [1111] [1111] [1111] [1111] [1111] [1111] [1111] [1111] [1111] [1111] [1111] [1111] [1111] [1111] [1111] [1111] [1111] [1111] [1111] [1111] [1111] [1111] [1111] [1111] [1111] [1111] [1111] [1111] [1111] [1111] [1111] [1111] [1111] [1111] [1111] [1111] [1111] [1111] [1111] [1111] [1111] [1111] [1111] [1111] [1111] [1111] [1111] [1111] [1111] [1111] [1111] [1111] [1111] [1111] [1111] [1111] [1111] [1111] [1111] [1111] [1111] [1111] [1111] [1111] [1111] [1111] [1111] [1111] [1111] [1111] [1111] [1111] [1111] [1111] [1111] [1111] [1111] [1111] [1111] [1111] [1111] [1111] [1111] [1111] [1111] [1111] [1111] [1111] [1111] [1111] [111] [1111] [1111] [1111] [1111] [1111] [1111] [1111] [1111] [1111] [1111] [1111] [1111] [1111] [1111] [1111] [1111] [1111] [1111] [1111] [1111] [1111] [1111] [1111] [1111] [1111] [1111] [1111] [1111] [1111] [1111] [1111] [1111] [1111] [1111] [1111] [1111] [1111] [1111] [1111] [1111] [1111] [1111] [1111] [1111] [1111] [1111] [1111] [1111] [1111] [1111] [1111] [1111] [1111] [1111] [111] [1111] [1111] [1111] [1111] [1111] [1111] [1111] [1111] [1111] [1111] [1111] [1111] [1111] [1111] [1111] [1111] [1111] [1111] [1111] [1111] [1111] [1111] [1111] [1111] [1111] [1111] [1111] [1111] [1111] [1111] [1111] [1111] [1111] [1111] [1111] [1111] [1111] [1111] [1111] [1111] [1111] [1111] [1111] [1111] [1111] [1111] [1111] [1111] [1111] [1111] [1111] [1111] [1111] [1111] [111] [1111] [1111] [1111] [1111] [1111] [1111] [1111] [1111] [1111] [1111] [1111] [1111] [1111] [1111] [1111] [1111] [1111] [1111] [1111] [1111] [1111] [1111] [1111] [1111] [1111] [1111] [1111] [1111] [1111] [1111] [1111] [1111] [1111] [1111] [1111] [1111] [111] [1111] [1111] [1111] [1111] [1111] [1111] [1111] [1111] [111] [1111] [1111] [1111] [1111] [1111] [1111] [1111] [1111] [1111] [1111] [1111] [1111] [1111] [1111] [1111] [1111] [1111] [1111] [1111] [1111] [1111] [1111] [1111] [1111] [1111] [1111] [1111] [
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         140
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   100
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    294
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        141 LeuLeuSerProArgPro11eSerTyrLeuLysGlySerSerGlyGlyProLeuLeuCys 160
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             174
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        234
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  80
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            54
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 40
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  09
                                                                                                                                                                                                                                                                                                                                                                                                  Modified hepatitis C virus (HCV) NS3 protease comprising at least 1 substitution of a hydrophobic alpha-helix 0 amino acid to a hydrophilic amino acid, useful for screening inhibitors that may treat hepatitis C
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 61 ThrSerIleAsnGlyValLeuTrpThrValTyrHisGlyAlaGlyThrArgThrIleAla
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              21 TyralaglnGlnThrArgGlyGluGlnGlyThrGlnLysThrSerHisThrGlyArgAsp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  41 LysasnGlnValGluGlyGluValGlnIleValSerThrAlaThrGlnThrPheLeuAla
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      175 ACCIGCATCAACGGIGITIGCIGGACCGITIACCACGGIGCIGGIACCCGIACCAICGCI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SerProLysGlyProValThrGlnMetTyrThrAsnValAspLysAspLeuValGlyTrp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    121 LeuValThrArgHisAlaAspValIleProValArgArgArgGlyAspSerArgGlySer
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          588
183
29
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Sequence 588 BP; 103 A; 180 C; 156 G; 149 T; 0 other;
                                                                                                                                                                                >
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Length:
Matches:
Conservative:
Mismatches:
                                                                                                                                                                                     Goldfarb
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                US-09-965-594-22 (1-197) x AAA73329 (1-588)
                                                                                                                                                                           Zhang Y,
                                                                                               ဗ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Claim 26; Fig 12; 66pp; English.
                                                                                          SQUIBB
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     5.38e-78
929.00
94.42%
92.89%
91.44%
                                                                                                                                                                                Weinheimer S,
99US-0115271
                                                                                          (BRIM ) BRISTOL-MYERS
                                                                                                                                                                                                                                                                      2000-465976/40.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Percent Similarity:
Best Local Similarity:
                                                                                                                                                                                                                                                                                                                      P-PSDB; AAB15220
                                                                                                                                                                                Wittekind M,
08-JAN-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Alignment Scores:
Pred. No.:
```

40 54

9

174

80

234

g

ò

a

ö

```
101 GlnAlaProGlnGlySerArgSerLeuThrProCysThrCysGlySerSerAspLeuTyr 120
                                                                                                                                                                                                                                                                                                                                                                                      LeuValThrArgHisAlaAspValIleProValArgArgArgGlyAspSerArgGlySer 140
                                                                                                                                                                                                                                                                                                                                                                                                                                                        LeuLeuSerProArgProIleSerTyrLeuLysGlySerSerGlyGlyProLeuLeuCys
                                                                                                                                                                                                                                 MetLysLysLysGlySerValVallleValGlyArglleAsnLeuSerGlyAspThrAla
                                                  41 LysAsnGlnValGluGlyGluValGlnIleValSerThrAlaThrGlnThrPheLeuAla
                                                                                                                                                                                21 TyralacinginThrargclyGluGinGlyThrGlnLysThrSerHisThrGlyArgasp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    181 AlaValAspPheIleProValGluSerLeuGluThrThrMetArgSerPro 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       US-09-965-594-22 (1-197) x AAA73328 (1-588)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Chimeric BVDV/HCV NS3-wt sequence.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ABA95615 standard; DNA; 12734 BP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             99US-0344456
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             990S-0344456
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Hong 2, Lai VCH, Lau JYN;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              (SCHE ) SCHERING CORP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               WPI; 2002-121103/16.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           US6326137-BI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             25-JUN-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             25-JUN-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        21-MAR-2002
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          04-DEC-2001
                                                                                                                                                                                                   115
                                                                                                                                                                                                                                                                  175
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      141
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     535
                                                                                                                                                                                                                                   61
                                                                                                                                                                                                                                                                                                                                                                                                                                     121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                    355
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      161
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    475
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ABA95615;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        RESULT 9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ABA9561
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Dp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                            qq
                                                                                                                                                                  ö
                                                                                                                                                                                           qq
                                                                                                                                                                                                                        ð
                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                         ď
                                                                                                                                                                                                                                                                                                                          쉽
                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    à
                                  ò
                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  The present sequence is the coding sequence for a fusion protein created using the Hepatitis C virus (HCV) NS3 and NS4A protease enzymes. These proteins are both essential for the replication of the virus, acting to cleave its replicative proteins from the polyprotein produced from the HCV genome. Inhibitors of the two proteins should be effective as antiviral treatments of HCV infection. This is useful as HCV can lead to chronic liver disease such as cirrhosis. Ilver failure and liver cancer. The present invention concerns a number of NS3 mutants and NS3-NS4A fusion proteins which can be used to identify inhibitors of this type, as well as enabling structural studies of the protease and
161 ProAlaClyHisAlaValGlyIlcPheArgAlaAlaValSerThrArgGlyValAlaLys 180
                   Modified hepatitis C virus (HCV) NS3 protease comprising at least 1 substitution of a hydrophobic alpha-helix 0 amino acid to a hydrophilic amino acid, useful for screening inhibitors that may treat hepatitis C
                                                                                                                                                                                                                                                                                              Hepatitis; NS3 protease; viral replication; chronic liver disease;
liver failure; liver cancer; ds.
                                                                Hepatitis C virus NS4A-NS3 fusion protease coding sequence #1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              588
181
1
13
13
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Sequence 588 BP; 97 A; 183 C; 153 G; 155 T; 0 other;
                                                                                                                                                                                                                                                                                                                                                                                                                                                   /product- "NS3-NS4A fusion protein"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Goldfarb V;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Conservative:
Mismatches:
Indels:
Gaps:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Length:
Matches:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Wittekind M, Weinheimer S, Zhang Y,
                                                                                                                                                                                                                                                                                                                                                                                           Location/Qualifiers
1..588
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Disclosure; Fig 10; 66pp; English.
                                                                                                                                                                      AAA73328 standard; DNA; 588 BP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             protease:inhibitor complexes
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (BRIM ) BRISTOL-MYERS SQUIBB
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1.81e-75
902.00
92.39%
91.88%
88.78%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      99US-0115271
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     2000WO-US00345
                                                                                                                                                                                                                                   19-DEC-2000 (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                     /*tag=
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      WPI; 2000-465976/40.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Percent Similarity:
Best Local Similarity:
Query Match:
DB:
                                                                                                                                                                                                                                                                                                                                                 Hepatitis C virus.
Synthetic.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      P-PSDB; AAB15212.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      WO200040707-A1
```

06-JAN-2000; 08-JAN-1999;

13-JUL-2000

AAA73328

160

```
Nucleic acid construct encoding chimeric Hepatitis C Virus (HCV)
                                           Pestivirus; Npro; protease; NS3; screening; ds.
                                                      Chimeric - Bovine viral diarrhea virus.
Chimeric - Hepatitis C virus.
```

Alignment Scores:

œ

AAX80355;

```
892
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            The present invention relates to a nucleic acid construct encoding a chimeric Hepatitis C virus (HCV)-pestivirus genome. The construct comprises a pestivirus genome where a Npro pestivirus protease gene is replaced with a gene encoding a functional HCV NS3 protease. Furthermore, each junction site recognised by the Npro protease is replaced with a junction site recognised by the HCV NS3 protease. The construct is useful for screening compounds that inhibit HCV in vivo by inhibiting HCV protease, where screening may be in cell culture or in an animal model. The present sequence is a chimmeric clone of BVDV (bovine viral diarrhea virus)/HCV NS3-wt, which was used to illustrate the present invention.
                                                                                                                                                                                                                                                                                                                                                                                                                                                  532
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   AlaProGlnGlySerArgSerLeuThrProCysThrCysGlySerSerAspLeuTyrLeu 121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 LeuSerProArgProlleSerTyrLeuLysGlyScrSerGlyGlyProLeuLeuCysPro 161
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            AlaGlyHisAlaValGlyIlePheArgAlaAlaValSerThrArgGlyValAlaLysAla 181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            IGCATCAAIGGGGIAIGCIGGACIGICIACCACGGGGCCGGAACGAGGACCAICGCAICA 652
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ProLysGlyProValThrGlnMetTyrThrAsnValAspLysAspLeuValGlyTrpGln 101
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             41
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               81
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 pestlvirus genome where the Npro protease gene is replaced with NS3 protease gene, useful for in vivo screening of compounds which inhibit HCV infection
                                                                                                                                                                                                                                                                                                                                                                                                                                    GlySerValValIleValGlyArgIleAsnLeuSerGlyAsp-----ThrAlaTyr
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             AlaGlnGlnThrArgGlyGluGlnGlyThrGlnLysThrSerHisThrGlyArgAspLys
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        GCCCAGCAGACGAGAGGCCTCCTAGGGTGTAAGATCACCAGTCTGACTGGCCGGGACAAA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       AsnGlnValGluGlyGluValGlnIleValSerThrAlaThrGlnThrPheLeuAlaThr
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                SerIleAsnGlyValLeuTrpThrValTyrHisGlyAlaGlyThrArgThrIleAlaSer
                                                                                                                                                                                                                                                        Sequence 12734 BP; 4032 A; 2604 C; 3295 G; 2803 T; 0 other;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ValAspPheIleProValGluSerLeuGluThrThrMetArgSer 196
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Length:
Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                                      US-09-965-594-22 (1-197) x ABA95615 (1-12734)
                                                                                                                                                                                                                                                                                                                                                                           Gaps:
                                                                 Example 2; Columns 17-28; 20pp; English.
                                                                                                                                                                                                                                                                                                  5.54e-72
882.50
92.31%
90.26%
86.86%
                                                                                                                                                                                                                                                                                                                          Percent Similarity:
Best Local Similarity:
Query Match:
DB:
                                                                                                                                                                                                                                                                                        Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            593
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    142
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              162
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          893
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     182
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   953
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             22
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         473
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       42
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   533
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               62
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          82
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    653
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 102
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               713
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                RESULT 10
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              AAX80355
ID AAX8
XX
                                                                                                                                                                                                                                                                                                                      Score:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 용
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        å
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   οÿ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     5
      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ŏ
```

AAX80355 standard; cDNA; 1998 BP

```
243
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            123
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   183
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          81
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     22 AlaGlnGlnThrArgGlyGluGlnGlyThrGlnLysThrSerHisThrGlyArgAspLys 41
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         NS4A-NS3 complex computating a central hydrophobic domain of native NCV NS4A-NS3 complex complex and an HCV NS3 serine protease domain, where the hydrophobic domain of native HCV NS4A serine protease domain, where the hydrophobic domain of native HCV NS4A serine protease domain, where the hydrophobic domain of native HCV NS4A serine protease domain. The present sequence encodes an example of the above complex. The covalent NS4A-NS3 complexes are useful for structural determination and determination of mode of binding of HCV inhibitors by NNR spectroscopy. They can also be used for detecting inhibitors of the protease activity and the Arpase activity of NS3. The covalent NS4A-NS3 complexes are more soluble, stable and active than the noncovalent protease-peptide complexes previously available.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            SerIleAsnGlyValLeuTrpThrValTyrHisGlyAlaGlyThrArgThrIleAlaSer
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  42 AsnGlnValGluGlyGluValGlnIleValSerThrAlaThrGlnThrPheLeuAlaThr
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            GlySerValValIleValGlyArgIleAsnLeuSerGlyAsp-----ThrAlaTyr
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 The present invention describes a covalent hepatitis C virus (HCV)
                                                                                   HCV; hepatitis C virus; single chain recombinant complex; linker; NS3 protease; NS4A cofactor; NS4A-NS3 complex; serine protease; hydrophobic domain; covalent complex; detection; inhibitor; ss.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1998
167
15
11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Sequence 1998 BP; 411 A; 595 C; 569 G; 423 T; 0 other;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Conservative:
Mismatches:
Indels:
                                                          HCV NS4A-NS3 complex encoding cDNA SEQ ID NO:105
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Length:
Matches:
                                                                                                                                                                                                                                                                                                                                                             ż
                                                                                                                                                                                                                                                                                                                                                                                                                                                   Disclosure; Page 166-169; 211pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                        New hepatitis C virus covalent complexes
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Gaps:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              US-09-965-594-22 (1-197) x AAX80355 (1-1998)
                                                                                                                                                                                                                                                                                                                                                             Weber PC,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      2.5e-72
875.50
92.86%
85.20%
86.17%
                                                                                                                                                                                                                                                        98WO-US24528
                                                                                                                                                                                                                                                                                    98US-0094331
                                                                                                                                                                                                                                                                                                   97US-0067315
                              (first entry)
                                                                                                                                                                                                                                                                                                                                                             Malcolm BA, Taremi SS,
                                                                                                                                                                                                                                                                                                                                 (SCHE ) SCHERING CORP
                                                                                                                                                                                                                                                                                                                                                                                          WPI; 1999-385385/32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Percent Similarity:
Best Local Similarity:
                                                                                                                                                 Hepatitis C virus.
                                                                                                                                                                                                                                                      24-NOV-1998;
                                                                                                                                                                                                                                                                                   28-JUL-1998;
28-NOV-1997;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Alignment Scores:
                                                                                                                                                                                            W09928482-A2
                              07-SEP-1999
                                                                                                                                                                                                                          10-JUN-1999
                                                                                                                                                                 Synthetic
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Ŋ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              62
a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
```

other;

569 G; 423 T; 0

ö

```
Sequence 1998 BP;
                                                   Percent Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  W09928482-A2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               07-SEP-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       10-JUN-1999
                      Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          AAX80354;
                                                                                                                                                                                                                                                     62
                                                                                                                                                                                                                                                                                                                                                                              122
                                                                                                                                                                                                                                                                                                                                                                                                 424
                                                                                                                                                                                                                                                                                                                                                                                                                      142
                                                                                                                                                                                                                                                                                                                                                                                                                                           184
                                                                        Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  RESULT 12
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           AAX80354
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                             g
                                                                                                                                                                   ò
                                                                                                                                                                                                                          q
                                                                                                                                                                                                                                                  δ
                                                                                                                                                                                                                                                                    a
                                                                                                                                                                                                                                                                                       õ
                                                                                                                                                                                                                                                                                                               qq
                                                                                                                                                                                                                                                                                                                                      à
                                                                                                                                                                                                                                                                                                                                                      a
                                                                                                                                                                                                                                                                                                                                                                            δ
                                                                                                                                                                                                                                                                                                                                                                                            q
                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                      q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                 δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                           ò
                                                                                                                                                                                      g
                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          NS4A-NS3 complex comprising a central hydrophobic domain of native HCV NS4A peptide, a linker, and an HCV NS3 serine protease domain, where the hydrophobic domain of native HCV NS4A peptide is tethered by the linker to the amino terminus of the HCV NS3 protease domain. The present sequence encodes an example of the above complex. The covalent NS4A-NS3 complexes are useful for structural determination and determination of mode of binding of HCV inhibitors by NHR spectroscopy. They can also be used for detecting inhibitors of the protease activity and the APPase activity of NS3. The covalent NS4A-NS3 complexes are more soluble, stable and active than the noncovalent protease-peptide complexes previously available.
                                                                                             543
                                                                                                                                                                   121
                                                    122 ValThrargHisAlaAspValIleProValArgArgArgGlyAspSerArgGlySerLeu 141
                                                                                                                          LeuSerProArgProIleSerTyrLeuLysGlySerSerGlyGlyProLeuLeuCysPro 161
 ProLysGlyProValThrGlnMetTyrThrAsnValAspLysAspLeuValGlyTrpGln 101
                                                                                                                                     102 AlaProGlnGlySerArgSerLeuThrProCysThrCysGlySerSerAspLeuTyrLeu
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    present invention describes a covalent hepatitis C virus (HCV)
                                                                                                                                                                                                                                                                                                                                                         HCV; hepatitis C virus; single chain recombinant complex; linker; NS3 protease; NS4A cofactor; NS4A-NS3 complex; serine protease, hydrophobic domain; covalent complex; detection; inhibitor; ss.
                                                                                                                                                                                                            ValAspPheIleProValGluSerLeuGluThrThrMetArgSerPro 197
                                                                                                                                                                                                                       604 GIGGACTTIGGCCCGFAGAGTCCATGGAAACTACTATGCGGTCTCCG 651
                                                                                                                                                                                                                                                                                                                                      HCV NS4A-NS3 complex encoding cDNA SEQ ID NO:109
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ż
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Disclosure; Page 179-182; 211pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           New hepatitis C virus covalent complexes
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Yao
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  PC,
                                                                                                                                                                                                                                                                         BP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Weber
                                                                                                                                                                                                                                                                         AAX80359 standard; cDNA: 1998
                                                                                                                                                                                                                                                                                                                                                                                                                                                                           98WO-US24528
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               98US-0094331
97US-0067315
                                                                                                                                                                                                                                                                                                                  (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Malcolm BA, Taremi SS,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (SCHE ) SCHERING CORP.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      WPI; 1999-385385/32.
                                                                                                                                                                                                                                                                                                                                                                                                 Hepatitis C virus.
Synthetic.
                                                                                                                                                                                                                                                                                                                                                                                                                                 W09928482-A2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                           24-NOV-1998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                28-JUL-1998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        28-NOV-1997;
                                                                                                                                                                                                                                                                                                                  07-SEP-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                      10-JUN-1999
                                                                                                                            142
                                                                                                                                                                                                             182
                                                                                                                                                                                                                                                                                              AAX80359;
                                                                                                                                                                                                                                                      RESULT 11
                                                                                                                                                                                                                                                              8
                                                                                                                                                                                      a
                  g
                                        õ
                                                         a
                                                                                  ò
                                                                                                                           ö
                                                                                                                                             셤
                                                                                                                                                                õ
                                                                                                                                                                                                              à
                                                                                                                                                                                                                            음
```

```
243
                                                                                                                                                                                                                                                                                                                                              303
                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ValThrArgHisAlaAspValIleProValArgArgArgGlyAspSerArgGlySerLeu 141
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             GTCACGAGACATGCTGACGTCATTCCGGTGCGCGGGGGGGACAGTAGGGGGGGAGCCTG 483
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 161
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             543
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     123
                                                                                                                                                                                                                                      183
                                                                                                                                                                                                                                                                61
                                                                                                                                                                                                                                                                                                                    81
                                                                                                                                                           21
                                                                                                                                                                                                              47
                                                                                                                                                                                                            LeuSerProArgProlleSerTyrLeuLysGlySerSerGlyGlyProLeuLeuCysPro
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  42 AsnGlnValGluGlyGluValGlnIleValSerThrAlaThrGlnThrPheLeuAlaThr
                                                                                                                                                                                                                                                                                                                    SerlleAsnGlyValLeuTrpThrValTyrHisGlyAlaGlyThrArgThrIleAlaSer
                                                                                                                                                          GlySerValValIleValGlyArgIleAsnLeuSerGlyAsp-----ThrAlaTyr
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                BCV; hepatitis C virus; single chain recombinant complex; linker; NS3 protease; NS4A cofactor; NS4A-NS3 complex; serine protease; hydrophobic domain; covalent complex; detection; inhibitor; ss.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           604 GIGGACTITGTGCCCGTAGAGICCAIGGAAACIACTAIGCGGICICCG 651
                                    1998
166
16
11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         HCV NS4A-NS3 complex encoding cDNA SEQ ID NO:104
                                                               Conservative:
Mismatches:
Indels:
                                                  Matches:
                                       Length:
                                                                                                       Gaps:
                                                                                                                               (1-1998)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         AAX80354 standard; cDNA; 1998 BP.
                                                                                                                                 US-09-965-594-22 (1-197) x AAX80359
411 A; 595
                                 4.77e-72
872.50
92.86%
84.69%
85.88%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Hepatitis C virus
Synthetic.
                                                                             Best Local Similarity:
```

```
544 TCGGGCACGCTGTGGGCATCTTCCGGCTGCCGTATGCACCCGGGGGTTGCGAAGGC 603
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                22 AlaGlnGlnThrArgGlyGluGlnGlyThrGlnLysThrSerHisThrGlyArgAspLys 41
                                                                                                                                                                          07-SEP-1999
                                                                                                                                                                                                                                                                                                                         W09928482-A2
                                                                                                                                                                                                                                                                                                                                                                           24-NOV-1998;
                                                                                                                                                                                                                                                                                                                                                                                                      28-JUL-1998;
                                                                                                                                                                                                                                                                                                                                                                                                                  28-NOV-1997;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                  10-JUN-1999
                                                                                                                                                                                                                                                                                                Synthetic.
                                                                                                                                                 AAX80345;
                                                                                            RESULT 13
                                                                                                          AAX80345
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Score:
                                                                                                                                                 a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                244 TGCGTCAACGGCGTGTGTGGACCGTTTACCATGGTGCTGGCTCAAAGACCTTAGCCGGC 303
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ValThrArgHisAlaAspValIleProValArgArgArgGlyAspSerArgGlySerLeu 141
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    LeuSerproArgProIleSerTyrLeuLysGlySerSerGlyGlyProLeuLeuCysPro 161
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ProLysGlyProValThrGlnMetTyrThrAsnValAspLysAspLeuValGlyTrpGln 101
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  AlaProGlnGlySerArgSerLeuThrProCysThrCysGlySerSerAspLeuTyrLeu 121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             162 AlaGlyHisAlaValGlyIlePheArgAlaAlaValSerThrArgGlyValAlaLysAla 181
                                                                                                                                                                                                            NS4A-NG3 complex comprising a central hydrophobic domain of native HCV NS4A-PD51cde, a linker, and an HCV NS3 serine protease domain, where the hydrophobic domain of native HCV NS4A peptide is tethered by the linker to the amino terminus of the HCV NS3 protease domain. The present sequence encodes an example of the above complex. The covalent NS4A-NS3 complexes are useful for structural determination and determination of mode of binding of HCV inhibitors by NMR spectroscopy. They can also be used for detecting inhibitors of the protease activity, when the helicase activity and the ATPASE activity of NS3. The covalent was the helicase activity of NS3. The covalent was the model of the ATPASE activity of NS3. The covalent was the model of the ATPASE activity of NS3. The covalent was the model of the ATPASE activity of NS3. The covalent was the model of the ATPASE activity of NS3. The covalent was the model of the ATPASE activity of NS3. The covalent was the atpase activity.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       41
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           81
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  GlySerValValIleValGlyArgIleAsnLeuSerGlyAsp-----ThrAlaTyr 21
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    AlaGlnGlnThrArgGlyGluGlnGlyThrGlnLysThrSerHisThrGlyArgAspLys
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        AsnGlnValGluGlyGluValGlnIleValSerThrAlaThrGlnThrPheLeuAlaThr
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SerIleAsnGlyValLeuTrpThrValTyrHisGlyAlaGlyThrArgThrIleAlaSer
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CTCTCCCCCAGGCCTGTCTCCTACTTGAAGGGCTCTTCGGGTGGTCCACTGCTCTGCCCT
                                                                                                                                                                                                                                                                                                                                    NS4A-NS3 complexes are more soluble, stable and active than the non-
                                                                                                                                                                                                   invention describes a covalent hepatitis C virus (HCV)
                                                                                                                                                                                                                                                                                                                                                 covalent protease-peptide complexes previously available
                                                                                                                                                                                                                                                                                                                                                                          Sequence 1998 BP; 410 A; 596 C; 568 G; 424 T; 0 other;
                                                                                                                                                                                                                                                                                                                                                                                                               1998
167
14
12
3
                                                                                                                                                                                                                                                                                                                                                                                                                 Length:
Matches:
Conservative:
Mismatches:
                                                                                              ż
                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Indels:
                                                                                                                                                                      Disclosure; Page 163-166; 211pp; English.
                                                                                              Yao
                                                                                                                                                 New hepatitis C virus covalent complexes
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Gaps:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        US-09-965-594-22 (1-197) x AAX80354 (1-1998)
                                                                                              Weber PC,
 98WO-US24528
                           98US-0094331.
97US-0067315.
                                                                                                                                                                                                                                                                                                                                                                                                               .91e-72
                                                                                                                                                                                                                                                                                                                                                                                                                          871.50
92.35$
85.20$
85.78$
                                                                                              Malcolm BA, Taremi SS,
                                                                  (SCHE ) SCHERING CORP
                                                                                                                     WPI; 1999-385385/32.
                                                                                                                                                                                                                                                                                                                                                                                                                                                       Best Local Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                          Percent Similarity:
24-NOV-1998;
                                                                                                                                                                                                                                                                                                                                                                                                   Alignment Scores:
                                                                                                                                                                                                     present
                           28-JUL-1998;
28-NOV-1997;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      22
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                102
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           364
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     122
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              424
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          142
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   484
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        42
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          62
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             82
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      304
                                                                                                                                                                                                                                                                                                                                                                                                                                                                Query Match:
DB:
                                                                                                                                                                                                                                                                                                                                                                                                                  Pred. No.:
 à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ô
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ជ្ជ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ö
```

```
64 GGTTCTGTTGTTGTTGGTAGAATTATTTTATCTGGTAGTGGTAGTATATCACGGCCTAC 123
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             5 GlySerValVallleValGlyArglleAsnLeuSerGlyAsp------ThrAlaTyr 21
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  The present invention describes a covalent hepatitis C virus (HCV) NS4ANS3 complex comprising a central hydrophobic domain of native HCV NS3 serine protease domain, where the hydrophobic domain of native HCV NS3 serine protease domain, where the hydrophobic domain of native HCV NS4A peptide is tethered by the linker to the amino terminus of the HCV NS3 protease domain. The present sequence encodes an example of the above complex. The covalent NS4A NS3 complexes are useful for structural determination and determination of mode of binding of HCV inhibitors by NMR spectroscopy. They can also be used for detecting inhibitors of the protease activity, the helicase activity and the ATPase activity of NS3. The covalent NS4A-NS3 complexes are more soluble, stable and active than the noncovalent protease-peptide complexes previously available.
                                                                                                                                                                                                                                                                                                                                                                                                                                                             HCV; hepatitis C virus; single chain recombinant complex; linker; NS3 protease; NS4A cofactor; NS4A-NS3 complex; serine protease; hydrophobic domain; covalent complex; detection; inhibitor; ss.
Sequence 651 BP; 120 A; 187 C; 200 G; 144 T; 0 other;
                                                                                                                                                                                                                                                                                                                                                                                                  HCV NS4A-NS3 complex encoding cDNA SEQ ID NO:95
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Conservative:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Length:
Matches:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ż
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Disclosure; Page 147-148; 211pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                New hepatitis C virus covalent complexes
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Yao
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        US-09-965-594-22 (1-197) x AAX80345 (1-651)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Weber PC,
                                                                                                                                                                                                    ВР
                                                                                                                                                                                                 AAX80345 standard; cDNA; 651
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              98US-0094331
97US-0067315
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                98WO-US24528
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    868.50
92.82%
85.13%
85.48%
                                                                                                                                                                                                                                                                                                                                    (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Malcolm BA, Taremi SS,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               (SCHE ) SCHERING CORP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               WPI; 1999-385385/32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Best Local Similarity:
Query Match:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Percent Similarity:
```

```
ValThrArgHisAlaAspValIleProValArgArgArgGlyAspSerArgGlySerLeu 141
                                                                                                                                                                                                                                                AlaGlyHisAlaValGlyIlePheArgAlaAlaValSerThrArgGlyValAlaLysAla 181
NS4A peptide, a linker, and an HCV NS3 serinc protease domain, where the hydrophobic domain of native HCV NS4A peptide is tethered by the linker to the amino terminus of the HCV NS3 protease domain. The present sequence encodes an example of the above complex. The covalent NS4A-NS3 complexes are useful for structural determination and determination of mode of binding of HCV inhibitors by NMR spectroscopy. They can also be used for detecting inhibitors of the protease activity, the helicase activity and the ATPASE activity of NS3. The covalent NS4A-NS3 complexes are more soluble, stable and active than the noncovalent protease-peptide complexes previously available.
                                                                                                                                                                                                                                                                                         102 AlaProGlnGlySerArgSerLeuThrProCysThrCysGlySerSerAspLeuTyrLeu
                                                                                                                                                                                                                                                                                22 AlaGlnGlnThrArgGlyGluGlnGlyThrGlnLysThrSerHisThrGlyArgAspLys
                                                                                                                                                                                                                                                                                                                      42 AsnGlnValGluGlyGluValGlnIleValSerThrAlaThrGlnThrPheLeuAlaThr
                                                                                                                                                                                                                                                                                                                                                                 SerIleAsnGlyValLeuTrpThrValTyrHisGlyAlaGlyThrArgThrIleAlaSer
                                                                                                                                                                                                                                        GlySerValVallleValGlyArglleAsnLeuSerGlyAsp-----ThrAlaTyr
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1998
166
15
12
                                                                                                                Sequence 1998 BP; 410 A; 596 C; 568 G; 424 T; 0 other;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CDNA SEQ ID NO:103.
                                                                                                                                             Length:
Matches:
Conservative:
Mismatches:
                                                                                                                                                                                       Indels:
                                                                                                                                                                                                                   US-09-965-594-22 (1-197) x AAX80358 (1-1998)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               BP.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             complex encoding
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              AAX80353 standard; cDNA; 1998
                                                                                                                                             1.13e-71
868.50
92.35%
84.69%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        (first entry)
                                                                                                                                                                          Best Local Similarity:
                                                                                                                                                                 Percent Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             HCV NS4A-NS3
                                                                                                                                    Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       07-SEP-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        162
                                                                                                                                                                                                                                                                                                                                                                 62
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       122
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           424
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 182
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    604
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 AAX80353;
                                                                                                                                                                                                                                       S
                                                                                                                                                                                                                                                                                                                                                                                                                           304
                                                                                                                                                                                      Query Match:
                                                                                                                                               ..
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        RESULT 15
                                                                                                                                                         Score:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               XEXEXEX
                                                                                                                                                                                                                                                                                                                                                                                                                         đ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                   셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
  ga
                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                              à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              à
                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                          a
                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                       à
                                                                                                                                   363
                                                                                                                                                                                                141
                                                                                                                                                                                                                  483
                                                                                                                                                                                                                                       161
                                                                                                                                                                                                                                                          543
                                                                                                                                                                                                                                                                               181
                                                                                                                                                                                                                                                                                                   603
  303
                                                                                                                101
                                                                                                                                                       121
                                                                                                                                                                           423
                                                   243
                               61
                                                                       81
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          The present invention describes a covalent hepatitis C virus (HCV) NS4A-NS3 complex comprising a central hydrophobic domain of native HCV
                                                                                                                          ProLysGlyProValThrGlnMctTyrThrAsnValAspLysAspLeuValGlyTrpGln
                                                                                                                                                                                                                                       LeuSerProArgProlleSerTyrLeuLysGlySerSerGlyGlyProLeuLeuCysPro
                                                                                                                                                                                                                                                 ValThrArgHisAlaAspValIlcProValArgArgArgGlyAspSerArgGlySerLeu
                                                                        SerIleAsnGlyValLeuTrpThrValTyrHisGlyAlaGlyThrArgThrIleAlaSer
                               AsnGlnValGluGlyGluValGlnIleValSerThrAlaThrGlnThrPheLeuAlaThr
                                                                                                                                                                                                                                                                                                                                                                                                                                                               HCV; hepatitis C virus; single chain recombinant complex; linker; NS3 protease; NS4A cofactor; NS4A-NS3 complex; serine protease; hydrophobic domain; covalent complex; detection; inhibitor; ss.
                                                                                                                                                                                                                                                                                                                      ValAspPheIleProValGluSerLeuGluThrThrMetArgSer 196
                                                                                                                                                                                                                                                                                                                                  604 GTGGACTTTGTGCCCGTAGAGTCCATGGAAACTACTATGCGGTCT 648
                                                                                                                                                                                                                                                                                                                                                                                                                                               HCV NS4A-NS3 complex encoding cDNA SEQ ID NO:108
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ż
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Disclosure; Page 176-179; 211pp; English
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 hepatitis C virus covalent complexes
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Yao
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        PC,
                                                                                                                                                                                                                                                                                                                                                                                    BP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Weber
                                                                                                                                                                                                                                                                                                                                                                                    CDNA; 1998
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 98WO-US24528
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    98US-0094331
97US-0067315
                                                                                                                                                                                                                                                                                                                                                                                                                           (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Taremi SS,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (SCHE ) SCHERING CORP.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             WPI; 1999-385385/32
                                                                                                                                                                                                                                                                                                                                                                                   AAX80358 standard;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    28-JUL-1998;
28-NOV-1997;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        W09928482-A2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 24-NOV-1998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        BA,
                                                                                                                                                                                                                                                                                                                                                                                                                           07-SEP-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            10-JUN-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Synthetic.
                                                                                                                                                                                                                                                                                                                      182
                                                                       62
                                                                                           244
                                                                                                                82
                                                                                                                                  304
                                                                                                                                                       102
                                                                                                                                                                           364
                                                                                                                                                                                                                  424
                                                                                                                                                                                                                                       142
                                                                                                                                                                                                                                                         484
                                                                                                                                                                                                                                                                                                                                                                                                      AAX80358;
                               45
                                                                                                                                                                                               122
          124
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Malcolm
                                                                                                                                                                                                                                                                                                                                                              셤
          a
                               δ
                                                g
                                                                ογ
                                                                                      a
                                                                                                               ò
                                                                                                                              a
                                                                                                                                                     õ
                                                                                                                                                                    셤
                                                                                                                                                                                         οy
                                                                                                                                                                                                              q
                                                                                                                                                                                                                                    ç
                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                 q
                                                                                                                                                                                                                                                                                                                      ò
```

61

41

243

81

483

```
The present invention describes a covalent hepatitis C virus (HCV) NSAA-NS3 complex comprising a central hydrophobic domain of native HCV NSAA peptide, a linker, and an HCV NS3 serine protease domain, where the hydrophobic domain of native HCV NSAA peptide is tethered by the linker to the amino terminus of the HCV NSAA peptide is tethered by the linker to the amino terminus of the HCV NS3A protease domain. The present sequence encodes an example of the above complex. The covalent NSAA-NS3 complexes are useful for structural determination and determination of mode of binding of HCV inhibitors by NNR spectroscopy. They can also be used for detecting inhibitors of the protease activity, the helicase activity and the ATPASE activity of NS3. The covalent NSAA-NS3 complexes are more soluble, stable and active than the noncovalent protease-peptide complexes previously available.
HCV; hepatitis C virus; single chain recombinant complex; linker; NS3 protease; NS4A cofactor; NS4A-NS3 complex; serine protease; hydrophobic domain; covalent complex; detection; inhibitor; ss.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Yao N;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Disclosure; Page 160-162; 211pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    New hepatitis C virus covalent complexes
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Weber PC,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            98WO-US24528
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          98US-0094331
97US-0067315
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Malcolm BA, Taremi SS,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                (SCHE ) SCHERING CORP.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           WPI; 1999-385385/32
                                                                                                                                                                                                          Hepatitis C virus
                                                                                                                                                                                                                                                                                                                                               W09928482-A2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     28-JUL-1998;
28-NOV-1997;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            24-NOV-1998;
                                                                                                                                                                                                                                                                                                                                                                                                                                        10-JUN-1999
                                                                                                                                                                                                                                                  Synthetic.
              SON COCCOCCON A STAN A
```

AlaProGlnGlySerArgSerLeuThrProCysThrCysGlySerSerAspLeuTyrLeu 121 243 22 AlaGlnGlnThrArgGlyGluGlnGlyThrGlnLySThrSerHisThrGlyArgAspLys 41 42 AsnGlnvalGluGlyGluValGlnIleValSerThrAlaThrGlnThrPheLcuAlaThr 61 62 SerlleAsnGlyValLeuTrpThrValTyrHisGlyAlaGlyThrArgThrIleAlaSer 81 Sequence 1998 BP; 410 A; 596 C; 568 G; 424 T; 0 other; 1998 166 15 12 Length:
Matches:
Conservative:
Mismatches:
Indels: Gaps: US-09-965-594-22 (1-197) x AAX80353 (1-1998) 1.4e-71 867.50 92.35% 84.69% 85.38% Percent Similarity: Best Local Similarity: Query Match: DB: Alignment Scores: a g ò đ qq õ g ŏ ò ò ò

23	1 9	83	61	43	81	03			
364 GCGCCCCCCGGGGCGCGTTCCTTGACACCATGCACCTGTGGCAGCTCAGACCTTTACTTG 423	122 ValThrArgHisAlaAspValIleProValArgArgArgGlyAspSerArgGlySerLeu 141	424 GTCACGAGACATGCTGACGTCATTCCGGTGCGCCGGGGGGGG	142 LeuSerProArgProlleSerTyrLeuLysGlySerSerGlyGlyProLeuLeuCysPro 161	484 CICICCCCCAGGCCTGTCCTCTTGAAGGCCTCTTCGGGTGGTCCACTGCTCTGCCCT 543	162 AlaGlyHisAlaValGlyIlePheArgAlaAlaValSerThrArgGlyValAlaLysAla 181	::	182 ValAspPheIleProValGluSerLeuGluThrThrMetArgSerPro 197	- 0	
364	122	424	142	484	162	544	182	604	
ορ	ογ	qq	Qy	qa	Qy	ą	oy	Ωp	

Search completed: August 30, 2003, 19:48:21 Job time : 188.939 secs

'n

```
Sequence 23, Application US/09965594

Sequence 23, Application US/09965594

Patent No. US20020016642A1

GENERAL INFORMATION:
APPLICANT: Wittekind, Michael
APPLICANT: Weinheamer, Steven
APPLICANT: Canage, Yaquen
TITLE OF INVENTION: Modified Forms of Hepatitis C NS3 Protease for
TITLE OF INVENTION: Modified Forms of Hepatitis C NS3 Protease for
TITLE OF INVENTION: Modified Forms of Protease: Inhibitor Complexes
FILE REFERENCE: D8178equences
CURRENT APPLICATION NUMBER: US/09/965,594
CURRENT APPLICATION NUMBER: 60/115,271
PRIOR APPLICATION NUMBER: 60/115,271
PRIOR FILING DATE: 1999-01-08
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 23
LENTH: 594
                                                                                                                                                            Sequence 4, Appli
Sequence 5, Appli
Sequence 1, Appli
Sequence 1, Appli
Sequence 1, Appli
Sequence 5, Appli
Sequence 5, Appli
Sequence 6, Appli
Sequence 1, Appli
Sequence 1, Appli
                                                                                                                                                                                                                                                                                                                                                                                                                                                                    19, Appl
2, Appl
1, Appl
1, Appl
120, Appl
17, Appl
17, Appl
13, Appl
13, Appl
16, Appl
16, Appl
16, Appl
16, Appl
16, Appl
17, Appl
18, Appl
18, Appl
16, Appl
16, Appl
17, Appl
18, Appl
18, Appl
19, Appl
11, A
                                                                                                                                                                                                                                                                                                                                                        Sequence 5, Sequence 5, Sequence 19, Sequence 19, Sequence 19, Sequence 19, Sequence 2, Sequence 12, Sequence 11, Sequence
                                                                                  Sequence
Sequence
Sequence
Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Sequence
Sequence
Sequence
Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Sequence Sequence Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Sequence
                                                              Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                               US-09-238-076-119
US-09-955-937-139
US-09-965-594-1
US-09-861-584-1
US-09-881-654-1
US-09-881-654-1
US-09-921-397-120
US-09-921-397-120
US-09-921-397-120
US-09-921-397-120
US-09-929-955-131
US-09-919-901-8
US-09-919-901-8
US-09-919-901-8
US-09-919-901-8
US-09-919-901-8
US-09-884-456-65
US-10-259-275-41
US-10-259-275-41
US-10-005-469-1
US-09-965-594-21

US-09-965-594-19

US-09-965-594-17

US-09-965-594-15

US-09-965-594-13

US-09-965-594-13

US-09-965-594-13

US-09-965-594-13

US-09-742-659-3
                                                                                                                                                                                                                                            US-09-238-076-1

US-09-995-937-1

US-09-917-563-1

US-09-238-076-5

US-09-995-937-5

US-09-916-359-1
     TYPE: DNA
CRCANISM: Hepatitis C virus
US-09-965-594-23
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    836
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10803
10
  841.5
841.5
841.5
831.5
831.5
831.5
831.5
821.5
821.5
821.5
821.5
821.5
821.5
821.5
821.5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  RESULT 1
                                                        Command line parameters:
-MODEL-frame+_p2n,model -DEV-xlp
-Q-/cgn2_1/USPTO_2801/US09965594/runat_29082003_151920_28367/app_query.fasta_11.2872
-Q-/cgn2_1/USPTO_spool/US09965594/runat_29082003_151920_28367/app_query.fasta_11.2872
-DB--UDB.ished_Applications_NA -OFMT-fastap -SUFFIX*npb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS-bits -START=1 -END=-1 -MATRIX*blosum62
-TRANS-human.40.cdi -LIST=45 -DOCALIGN=200 -THE_SCORE=pct -THE_MAX=100
-MAXLEN=0.000000000 -USER=US09965594_GCGN_1_1.864_Grunat_29082003_151920_28367
-NCPG-6 -LCPG-3 -NO_MAAP -LARGEQUERY -NG_SCORES=0 -WAIT -DSSELOCK=100
-EOMGLOG -DEV_INMEOUT=120 -WARN_TIMEOUT=30 -THREDS=1 -XGAPOPP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Sequence 23, Appl
                                                                                                                                                                                August 30, 2003, 19:26:03; Search time 176.482 Seconds (without alignments) 2560.981 Million cell updates/sec
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.
                                                                                                                                                                                                                                                                                                                                                     1 MKKKGSVVIVGRINLSGDTA.....VAKAVDFIPVESLETIMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Description
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       /cgn2_6/ptodata/2/pubpna/USO7_PUBCOMB.Seq:*
/cgn2_6/ptodata/2/pubpna/DCT_NEW_PUB.seq:*
/cgn2_6/ptodata/2/pubpna/DCT_NEW_PUB.seq:*
/cgn2_6/ptodata/2/pubpna/USO6_NEW_PUB.seq:*
/cgn2_6/ptodata/2/pubpna/USO6_NEW_PUB.seq:*
/cgn2_6/ptodata/2/pubpna/USO8_NEW_PUB.seq:*
/cgn2_6/ptodata/2/pubpna/USO8_NEW_PUB.seq:*
/cgn2_6/ptodata/2/pubpna/USO8_NEW_PUB.seq:*
/cgn2_6/ptodata/2/pubpna/USO8_NEW_PUB.seq:*
/cgn2_6/ptodata/2/pubpna/USO9_NEW_PUB.seq:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*/cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*
                         GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.
                                                                                                                                     nucleic search, using frame_plus_p2n model
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             US-09-965-594-23
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Total number of hits satisfying chosen parameters:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1533700 segs, 1147125425 residues
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SUMMARIES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Published_Applications_NA:*
                                                                                                                                                                                                                                                                                                                                                                                                                                     0.5
                                                                                                                                                                                                                                                                                                                                                                                                BLOSUM62
Xgapop 10.0 , Xgapext 0
Ygapop 10.0 , Ygapext (
Fgapop 6.0 , Fgapext
helop 6.0 , Delext
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Minimum DB seq length: 0
Maximum DB seq length: 200000000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             594 10
                                                                                                                                                                                                                                                                                                  US-09-965-594-22
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Query
Match Length DB
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             100.0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Score
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1016
                                                                                                                                                                                                                                                                                                                                                                                                              Scoring table:
                                                                                                                                                                                                                                                                                                                              Perfect score:
                                                                                                                                        OM protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Database :
                                                                                                                                                                                                                                                                                                                                                           Sequence:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            .
¥
                                                                                                                                                                                           Run on:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Result
```

a

```
US-09-965-594-19
                  Alignment Scores
JS-09-965-594-21
                                                           Query Match:
DB:
                           ..
⊗
                                  Score:
                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                         οχ
                                                                                                                                                                                                                                                                                                                                                             a
                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                              QQ
                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                 8
                                                                                                                    g
                                                                                                                                     ò
                                                                                                                                                    qq
                                                                                                                                                                                    g
                                                                                                                                                                                                                    qu
                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                          a
                                                                                                                                                                                                                                                                             ő
                                                                                                     ö
                                                                                                                                                                      ^{\circ}_{Y}
                                                                                                                                                                                                       ô
                                                                                                                                                                                                        240
                                                                                                                                                                                                                         120
                                                                                                    9
                                                                                                                     40
                                                                                                                                                                                        80
                                                                                    50
                                                                                                                                                                                                                                                                                                                                                                                                                                             Sequence 21, Application US/09965594

Fatent No. US20020106642A1

GENERAL INFORMATION:

APPLICANT: Wittekion:

APPLICANT: Wainheimer, Steven

APPLICANT: Wainheimer, Steven

APPLICANT: Coldfarb, Valentina

APPLICANT: Coldfarb, Valentina

TITLE OF INVENTION: Facilitating Inhibitor Screening and Structural Studi

TITLE OF INVENTION: Facilitating Inhibitor Complexes

TITLE OF INVENTION: OF Protease: Inhibitor Complexes

TILE REFERENCE: DB17Sequences

CURRENT FILING DATE: 2001-09-27

FRICE FILING DATE: 2001-09-27

PRIOR APPLICATION NUMBER: 60/115,271

PRIOR FILING DATE: 1999-01-08

NUMBER OF SEQ ID NOS: 26

SOFTWARE: Patentin Ver. 2.0

SEQ ID NO 21
                                                                                                                                                                                       MetLysLysLysClySerValValIleValGlyArg1leAsnLeuSerGlyAspThrAla
                                                                                                     1 ATGAAAAAAAAAAAGGATCCGTTGTTATCGTCGGCCGTATCAACCTGTCCGGTGACACCGCT
                                                                                                                     21 TyralaGlnGlnThrArgGlyGluGlnGlyThrGlnLysThrSerHisThrGlyArgAsp
                                                                                                                                                                                                                                                                                                                                                                                               594
197
0
0
0
      Length:
Matches:
Conservative:
Mismatches:
Indels:
                                                                   US-09-965-594-22 (1-197) x US-09-965-594-23 (1-594)
                                                    Gaps:
     4.57e-109
1016.00
100.00%
100.00%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   TYPE: DNA ORGANISM: Hepatitis C virus
                     Percent Similarity:
Best Local Similarity:
Query Match:
                                                                                                                                                                                                                                                                                                                                                                                                                                   RESULT 2
US-09-965-594-21
Alignment Scores:
Pred. No.:
                                                                                                                                                                                                    q
                                                                                                                                                                                                                         9.
O.
                                                                                                                                                                                                                                                           0.7
                                                                                                                                                                                                                                                                                                         면
장 연
                                                                                                                                                                                                                                                                                                                                                                              Q
                                                                                                                                                                                                                                                                                                                                                                                                             a
                                                                                                  g
                                                                                                                                 QQ
                                                                                                                                                                  g
                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                        õ
                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                    ö
                                                                                                                    Q
                                                                                                                                                ò
```

```
APPLICANT: Witherind, Michael
APPLICANT: Witherind, Michael
APPLICANT: Weinheimer, Steven
APPLICANT: Zhang, Yaqun
APPLICANT: Zhang, Yaqun
APPLICANT: Chang, Yaqun
TITLE OF INVENTION: Modified Forms of Hepatitis C NS3 Frotease for
TITLE OF INVENTION: Facilitating Inhibitor Screening and Structural Studies
TITLE OF INVENTION: of Protease:Inhibitor Complexes
FILE REFERENCE: DB17Sequences
CURRENT APPLICATION NUMBER: US/09/965,594
CURRENT FILING DATE: 2001-09-27
PRIOR FILING DATE: 1999-01-08
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    180
                                                                                                                                                                                                                                                                       240
                                                                                                                                                                                                                                                                                             100
                                                                                                                                                                                                                                                                                                                     300
                                                                                                                                                                                                                                                                                                                                            120
                                                                                                                                                                                                                                                                                                                                                                  360
                                                                                                                                                                                                                                                                                                                                                                                          140
                                                                                                                                                                                                                                                                                                                                                                                                               420
                                                                                                                                                                                                                                                                                                                                                                                                                                      160
                                                                                                                                                                                                                                                                                                                                                                                                                                                             480
                                                                                                                                                                                                  9
                                                                                                                                                                                                                                                 80
                                                                                                                              9
                                                                                                                                                                                                                                                                                                                                          141 LeuLeuSerProArgProlleSerTyrLeuLysGlySerSerGlyGlyProLeuLeuCys
                                                                                                                                                                                                                                                                                                                                                                                                                                                   41 LysAsnGlnValGluGlyGluValGlnIleValSerThrAlaThrGlnThrPheLeuAla
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                594
196
0
0
0
 Length:
Matches:
Conservative:
Mismatches:
Indels:
Gaps:
                                                                                 US-09-965-594-22 (1-197) x US-09-965-594-21 (1-594)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ; Sequence 19, Application US/09965594; Patent No. US20020106642A1; GENERAL INFORMATION:
2.28e-108
1010.00
99.49%
99.49%
99.41%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   NUMBER OF SEQ ID NOS: 26
SOFTWARE: Patentin Ver. 2.0
                      Percent Similarity:
Best Local Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SEQ ID NO 19
LENGTH: 594
```

```
06e-105
                                    virus
                                                                                 980.00
97.46%
96.45%
96.46%
                                    ပ
                        TYPE: DNA
CORGANISM: Hepatitis
US-09-965-594-17
                                                                                         Percent Similarity:
Best Local Similarity:
Query Match:
                                                               Alignment Scores:
Pred. No.:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  RESULT 5
US-09-965-594-15
      SEQ ID NO 17
LENGTH: 594
                                                                                                                                                                                                                                                                             61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                      161
                                                                                  Score:
                                                                                                                                                             ò
                                                                                                                                                                               a
                                                                                                                                                                                                  ö
                                                                                                                                                                                                                  qq
                                                                                                                                                                                                                                       οχ
                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                       õ
                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                  õ
                                                                                                                                                                                                                                                                                                                                                                       qq
                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                          a
                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                 õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       PP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Protease for and Structural Studies
                                                                                                                                                                                                                                                                                                                                                                       180
                                                                                                                                                                                                                                                                  GlnAlaProGlnGlySerArgSerLeuThrProCysThrCysGlySerSerAspLeuTyr 120
                                                                                                                                                                                                                                                                                                                                             LeuLeuSerProArgProIleSerTyrLeuLysGlySerSerGlyGlyProLeuLeuCys 160
                                                                                                                                                                                                                                                                                                                                                                                                                                   480
                                                                                                                                                                                                                     9
                                                                                                                                                                                                                                                          80
                                                                                                                                          50
                                                                                                                                                            9
                                                                                                                                                                              40
                                                                                                                                                                                                                    ThrSerIleAsnGlyValLeuTrpThrValTyrHisGlyAlaGlyThrArgThrIleAla
                                                                                                                                         MetLysLysLysGlySerValValIleValGlyArgIleAsnLeuSerGlyAspThrAla
                                                                                                                                                                             TyrAlaGlnGlnThrArgGlyGluGlnGlyThrGlnLySThrSerHisThrGlyArgAsp
                                                                                                                                                                                                                                                                                                                                                                                                                                  CTGCTGTCCCCGCGTCCCGATCTCCTACCTGAAAGGTTCCTCCGGTGGTCCCTGCTGTCC
                                                                                                                                                           ATGAAAAAAAAAGGATCCGTTGTTATCGTCGGCCGTATCAACCTGTCCGGTGACACCGCT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       APPLICANT: Wittekind, Michael
APPLICANT: Weinheimer, Steven
APPLICANT: Dang, Yaqun
APPLICANT: Johng, Yaqun
APPLICANT: Goldfarb, Valentina
TITLE OF INVENTION: Modified Forms of Hepatitis C NS3 Pro-
TITLE OF INVENTION: Facilitating Inhibitor Screening and
TITLE OF INVENTION: Of Protease:Inhibitor Complexes
FILE REFERENCE: DB17Sequences
CURRENT APPLICATION NUMBER: US/09/965,594
CURRENT FILING DATE: 2001-09-27
PRIOR FILING DATE: 1999-01-08
PRIOR FILING DATE: 1999-01-08
NUMBER OF SEQ ID NOS: 26
SOFTWARE: Patentin Ver. 2.0
                                                    594
22
00
0
                                                                      Conservative:
Mismatches:
Indels:
Gaps:
                                                                                                                       US-09-965-594-22 (1-197) x US-09-965-594-19 (1-594)
                                                     Length:
Matches:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Sequence 17, Application US/09965594
Patent No. US20020106642a1
GENERAL INFORMATION:
                                                  1.27e-106
995.00
98.98%
97.97%
              ; ORGANISM: Hepatitis C virus US-09-965-594-19
                                                                      Percent Similarity:
Best Local Similarity:
Query Match:
                                         Alignment Scores:
Pred. No.:
Score:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        09-965-594-17
                                                                                                                                                                              21
                                                                                                                                                                                                                                                                                                                                                                                          361
                                                                                                                                                                                                                                                        61
                                                                                                                                                                                                                                                                           181
                                                                                                                                                                                                                                                                                                                                   101
                                                                                                                                                                                                                                                                                                                                                                        121
                                                                                                                                                                                                                                                                                                                                                                                                               141
                                                                                                                                                                                                                                                                                                                                                                                                                                421
TYPE: 1
                                                                                                                                                         g
                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                    qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                        à
                                                                                                                                                                             õ
                                                                                                                                                                                           đ
                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                          QQ
                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                      a
                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ö
                                                                                                                                                                                                                à
                                                                                                                                                                                                                                                   δ
                                                                                                                                                                                                                                                                                                                                                                        οy
                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                   οy
```

```
Sequence 15, Application US/09965594
Patent No. US20020106642a1
GENERAL INFORMATION:
APPLICANT: Withekind, Michael
APPLICANT: Weinhelmer, Steven
APPLICANT: Weinhelmer, Steven
APPLICANT: Zhang, Yaqun
APPLICANT: Zhang, Yaqun
TILLE OF INVENTION: Modified Forms of Repatitis C NS3 Protease for
TITLE OF INVENTION: Pacilitating Inhibitor Screening and Structural Studies
TITLE OF INVENTION: of Protease:Inhibitor Complexes
TITLE OF INVENTION: Of Protease:Inhibitor Complexes
TITLE PEPERENCE: DB1/Sequences
CURRENT APPLICATION NUMBER: US/09/965,594
CURRENT APPLICATION NUMBER: 60/115,271
PRIOR APPLICATION NUMBER: 60/115,271
                                                                                                                                                                                      240
                                                                                                                                                                                                                                                                                                                                                             ProAlaGlyHisAlaValGlyIlePheArgAlaAlaValSerThrArgGlyValAlaLys 180
                                                                                                                                                            9
                                                                                                                                                                                                                                                                                                      ThrSerIleAsnGlyValLeuTrpThrValTyrHisGlyAlaGlyThrArgThrIleAla 80
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  181 ACCIGCAICAACGGTGTIGCIGGACCGTITACCACGGTGCTGGTACCCGTACCATCGCT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Conservative:
Mismatches:
Indels:
                                                                                                US-09-965-594-22 (1-197) x US-09-965-594-17 (1-594)
Length:
Matches:
```

```
21
                                                                                                                                                                                                                                                                                                                                                              81
                                                                                                            ..
Q
                                                                                                            Pred.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                       q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                              셤
                                                                                                                                                                                                                                                                                                                                                                                                   δ
                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                    ò
                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                         δ
                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                             οy
                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                       염
                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Sequence 25, Application US/09965594
Patent No. US20020106642a1
GENERAL INFORMATION:
APPLICANT: Witched, Michael
APPLICANT: Weinheimer, Steven
APPLICANT: Coldfach, Valentia
APPLICANT: APPLICATION: Modified Forms of Hepatitis C NS3 Frotease for
TITLE OF INVENTION: Pacilitating Inhibitor Screening and Structural Studies
TITLE OF INVENTION: of Protease:Inhibitor Complexes
FILE REFERENCE: DB17Sequences
CURRENT APPLICATION NUMBER: US/09/965,594
CURRENT-FILING DATE: 2001-09-27
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       100
                                                                                                                                                                                                                                                                                                                                                       LeuValThrargHisAlaaspvalIleProValArgArgArgArgGlyAspSerArgGlySer 140
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  LeuLeuSerProArgProIleSerTyrLeuLysGlySerSerGlyGlyProLeuLeuCys 160
                                                                                                                                                                                                                                  61 TACGCYCAGCAGACTCGAGGTGAGGAGGTTGCCAAGAAACCTCCCAGACGGTGGTGAC 120
                                                                                                                                                                                                                                                                                 180
                                                                                                                                                                                                                                                                                                                        240
                                                                                                                                                                                                                                                                                                                                                                                                                                            361 CTGGTTACCCGTCACGTGACGTTATCCCGGTTCGTCGTGGTGACTCCCGTGGTTCC 420
                                                                                                                                                                                                                                                                                                    80
                                                                                                                                                                                                                                                             9
                                                                                                                                                                                 50
                                                                                                                                                                                                   9
                                                                                                                                                                                                                     40
                                                                                                                                                                                                                                                                                                   SerProLysGlyProValThrGlnMetTyrThrAsnValAspLysAspLeuValGlyTrp
                                                                                                                                                                                TyralaGlnGlnThrArgGlyGluGlnGlyThrGlnLysThrSerHisThrGlyArgAsp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            594
187
2
0
0
                                                                                         Length:
Matches:
Conservative:
Mismatches:
Indels:
Gaps:
                                                                                                                                                              US-09-965-594-22 (1-197) x US-09-965-594-15 (1-594)
                                                                                        6.71e-103
963.00
95.94%
94.92%
94.78%
                                                    C virus
  26
; NUMBER OF SEQ ID NOS:
; SOFTWARE: Patentin Ver
; SEQ ID NO 15
; LENGTH: 594
; TYPE: DNA
; ORGANISM: Hepatitis C
US-09-965-594-15
                                                                                                          Percent Similarity:
Best Local Similarity:
Query Match:
DB:
                                                                                 Alignment Scores:
Pred. No.:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             US-09-965-594-25
                                                                                                                                                                                                                                                                                                                                                                                                                          121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                   141
                                                                                                                                                                                                                       21
                                                                                                                                                                                                                                                                                                                                            81
                                                                                                     Score:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        6 G
                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                   60 60
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                  ò
                                                                                                                                                                                                 q
                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                      dy
Oy
                                                                                                                                                                                                                                                                            a
                                                                                                                                                                                                                                                                                                o
G
                                                                                                                                                                                                                                                                                                                                       õ
                                                                                                                                                                                                                                                                                                                                                          q
                                                                                                                                                                                                                                                                                                                                                                            á
                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                         ö
```

```
Sequence 13, Application US/09965594
Patent No. US20020106642A1
GENERAL INFORMATION:
APPLICANT: Withekind, Michael
APPLICANT: Zhang, Yaqun
APPLICANT: Chang, Yaqun
TITLE OF INVENTION: Modified Forms of Hepatitis C NS3 Protease for TITLE OF INVENTION: Pacilitating Inhibitor Screening and Structural Studies FILE REFERENCE: DB17Sequences
FILE REFERENCE: DB17Sequences
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        420
                                                                                                                                                                                                                                                                                                                                                                                                                                    100
                                                                                                                                                                                                                                                                                                                                                                                                                                                           300
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           360
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  140
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 160
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        480
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       540
                                                                                                                                                                                                                                                                                                                                                                                                             240
                                                                                                                                                                                                                                                                                                                                                               180
                                                                                                                                                                                                                                                                                                                                       9
                                                                                                                                                                                                                                                                                                                                                                                      80
                                                                                                                                                                                                                                           20
                                                                                                                                                                                                                                                                 9
                                                                                                                                                                                                                                                                                         0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                101 GlnAlaProGlnGlySerArgSerLeuThrProCysThrCysGlySerSerAspLeuTyr
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 101 CCGGCTCCGCAGGGTTCCCGTTCCCTGACCCCGTGCACCTGCGGGTTCCTCCGACCTGTAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               161 ProAlaGlyHisAlaValGlyIlePheArgAlaAlaValSerThrArgGlyValAlaLys
                                                                                                                                                                                                                                          SerProLysGlyProValThrGlnMetTyrThrAsnValAspLysAspLeuValGlyTrp
                                                                                                                                                                                                                                                                                         TyrAlaGlnGlnThrArgGlyGluGlnGlyThrGlnLysThrSerHisThrGlyArgAsp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             594
185
0
0
0
                                                                                                                                                         Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                     US-09-965-594-22 (1-197) x US-09-965-594-25 (1-594)
                                                                                                                                   Length:
Matches:
                                                                                                                                                                                               Gaps:
PRIOR APPLICATION NUMBER: 60/115,271
PRIOR FILING DATE: 1999-01-08
NUMBER OF SEQ ID NOS: 26
SOFTWARE: PATENTIN VE. 2.0
SEQ ID NO 25
LENGTH: 594
TYPE: DNA
                                                                                                                               9.3e-100
936.00
93.91%
93.91%
92.13%
                                                                                     C virus
                                                                                   ; ORGANISM: Hepatitis
US-09-965-594-25
                                                                                                                                                        Percent Similarity:
Best Local Similarity:
Query Match:
DB:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         US-09-965-594-13
                                                                                                                        Alignment Scores:
```

S

```
Percent Similarity:
Best Local Similarity:
                                                                                                                                       Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            RESULT 9
US-10-133-133A-6
                                                                                      LENGTH: 588
TYPE: DNA
                                                                                                                                                                                                                                                                                                              41
                                                                                                                                                                                                                                                                                                                                                                     175
                                                                    SOFTWARE: E
                                                                                                                                                                                       Query Match
                                                                                                                                                ..
Q
                                                                                                                                                          Score:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    QQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  8
                                                                                                                                                                                                                                                     a
                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                   õ
                                                                                                                                                                                                                                                                                                                                                                                                         ΩD
                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                      ð
                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     APPLICANT: Wittekind, Michael
APPLICANT: Weinheimer, Steven
APPLICANT: Zhang, Yaqun
APPLICANT: Caldfarb, Valentina
TITLE OF INVENTION: Modified Forms of Hepatitis C NS3 Protease for
TITLE OF INVENTION: Facilitating Inhibitor Screening and Structural Studies
                                                                                                                                                                                                                                                                                                                                                                                                          534
                                                                                                                                                                                                                                                                                                                                                                     LeuValThrArgHisAlaAspValIleProValArgArgArgGlyAspSerArgGlySer 140
                                                                                                                                                                                                                                                                                                                                                                                                                                                           LeuLeuSerProArgProIleSerTyrLeuLysGlySerSerGlyGlyProLeuLeuCys 160
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           161 ProAlaGlyHisAlaValGlyIlePheArgAlaAlaValSerThrArgGlyValAlaLys 180
                                                                                                                                                                                                                                                                                          474
                                                                                                                                                                                                                                                                 234
                                                                                                                                                                                                                                                       40
                                                                                                                                                                                                                                                                                                                                80
                                                                                                                                                                                                                  20
                                                                                                                                                                                                                                    54
                                                                                                                                                                                                                                                                                                                               MetLysLysClySerValValIleValGlyArgIleAsnLeuSerGlyAspThrAla
                                                                                                                                                                                                                                                       TyralaGlnGlnThrArgGlyGluGlnGlyThrGlnLysThrSerH1sThrGlyArgAsp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   181 AlaValAspPheIleProValGluSerLeuGluThrThrMetArgSerPro 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            588
1183
9
1
                                                                                                                             Length:
Matches:
Conservative:
Mismatches:
                                                                                                                                                                                                US-09-965-594-22 (1-197) x US-09-965-594-13 (1-588)
                                                                                                                                                                   Indels:
                                                                                                                                                                            Gaps:
CURRENT APPLICATION NUMBER: US/09/965,594
CURRENT FILING DATE: 2001-09-27
PRIOR APPLICATION NUMBER: 60/115,271
PRIOR FILING DATE: 1999-01-08
NUMBER OF SEQ ID NOS: 26
SOFTWARE: PATENTIN Ver. 2.0
SEQ ID NO 13
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Sequence 4, Application US/09965594 Patent No. US20020106642Al GENERAL INFORMATION:
                                                                                                                           5.98e-99
929.00
94.42%
92.89%
                                                                                       ORGANISM: Hepatitis C virus
                                                                                                                                     Score:
Percent Similarity:
Best Local Similarity:
Query Match:
DB:
                                                                                                                    Alignment Scores:
                                                                                           ; OKGANISM: He|
US-09-965-594-13
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        RESULT 8
US-09-965-594-4
                                                                     LENGTH: 588
                                                                               TYPE: DNA
                                                                                                                                                                                                                                                       21
                                                                                                                                                                                                                                                                          55
                                                                                                                                                                                                                                                                                                              115
                                                                                                                                                                                                                                                                                                                                                                                                                                                                    355
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       141
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        415
                                                                                                                                                                                                                                                                                                                                                                                                                                                 121
                                                                                                                              Pred. No.:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                            õ
                                                                                                                                                                                                                                                                                                          q
                                                                                                                                                                                                                                                                                                                            ó
                                                                                                                                                                                                                                                                                                                                                qq
                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                     a
                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                             gg
                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                               Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    οχ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    d
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  òγ
                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                       õ
```

```
161 ProalaGlyHisalaValGlyIlePheargalaAlaValSerThrArgGlyValAlaLys 180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         81 SerProLysGlyProValThrGlnMetTyrThrAsnValAspLysAspLeuValGlyTrp 100
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         21 TyralacincinthrargciyclucinciythrcintysthrserHisthrciyargasp 40 [[[[[[[]]]]]]]]]]]
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          61 ThrSerlleAsnGlyValLeuTrpThrValTyrHisGlyAlaGlyThrArgThrIleAla
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          475 CCGCTGTCACGCTGTTGGTATCTTCCGTGCTGTTTGCACCCGTGGTGTTGCTAAA
                                                                                                                                                                                                                                                                                                                                                                                                    MetLysLysLysGlySerValVallleValGlyArglleAsnLeuSerGlyAspThrAla
                                                                                                                                                                                                                                                                                                                                                                                                                           LysasnGlnValGluGlyGluValGlnIleValSerThrAlaThrGlnThrPheLeuAla
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 APPLICANT: CATHERS, Brian
APPLICANT: NEUTEBON, SASKIA
APPLICANT: SHERAED, Michael
TITLE OF INVENTION: VIRAL ENZYME ACTIVATED PROTOTOXOPHORES
                                                                                                                                                                                                                                                       588
181
1
13
2
      Complexes
                                                                                                                                                                                                                                                                                      Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                          US-09-965-594-22 (1-197) x US-09-965-594-4 (1-588)
                                                                                                                                                                                                                                                       Length:
Matches:
TITLE OF INVENTION: of Protease:Inhibitor of FILE REFERENCE: DB17Sequences CURRENT APPLICATION NUMBER: US/09/965,594
CURRENT FILING DATE: 2001-09-27
PRIOR APPLICATION NUMBER: 60/115,271
PRIOR FILING DATE: 1999-01-08
NUMBER OF SEQ ID NOS: 26
SOFTWARE: Patentin Ver. 2.0
     of Protease: Inhibitor
                                                                                                                                                                                                                                                                                                                                            Gaps:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ; Sequence 6, Application US/10133133A
; Publication No. US20030114385A1
; GENERAL INFORMATION:
                                                                                                                                                                                                                                                     8.29e-96
902.00
92.39%
91.88%
88.78%
                                                                                                                                                                                        virus
                                                                                                                                                                                      ; ORGANISM: Hepatitis C
US-09-965-594-4
```

174

9

54

80

140

160

```
RESULT 11
US-09-238-076-1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   161
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          498
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           558
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 162 AlaGlyHisAlaValGlyIlePheArgAlaAlaValSerThrArgGlyValAlaLySAla 181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        GCCCAGCAGAACAAGGGGCCTCC1AGGGTGCATAATCACCAGCCTAACTGGCCGGGACAAA 138
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        41
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 78
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    142 LeuSerProArgProIleSerTyrLeuLysGlySerSerGlyGlyProLeuLeuCysPro
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    AlaGInGlnThrArgGlyGluGlnGlyThrGlnLySThrSerHisThrGlyArgAspLys
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     AsnGlnValGluGlyGluValGlnIleValSerThrAlaThrGlnThrPheLeuAlaThr
TITLE OF INVENTION: AND USE OF SAME TO TREAT VIRAL INFECTIONS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ValAspPheIleProValGluSerLeuGluThrThrMetArgSer 196
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        612
174
3
15
15
                                                                                                                                                                                                                                                                                                                                                                          Length:
Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     US-09-965-594-22 (1-197) x US-10-133-133A-6 (1-612)
          FILE REFERENCE: NB 2021.00
CURRENT APPLICATION NUMBER: US/10/133,133A
CURRENT FILING DATE: 2002.04-26
PRIOR APPLICATION NUMBER: 60/286,983
PRIOR FILING DATE: 2001-04-27
NUMBER OF SEQ ID NOS: 21
SEQ ID NO 6
SEQ ID NO 6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Sequence 3, Application US/09742659
Patent No. US20010034019A1
SERERAL INFORMATION:
APPLICANT: HONG, Zhi
APPLICANT: Zhong, Weldong
APPLICANT: Zhong, Weldong
APPLICANT: LIGHTANATONE
APPLICANT
                                                                                                                                                                                                                                                                                                                                                                        2.02e-91
864.50
90.77%
89.23%
85.09%
                                                                                                                                                                                                                                                                            ORGANISM: Hepatitus C. Virus
                                                                                                                                                                                                                                                                                                                                                                                                                            Percent Similarity:
Best Local Similarity:
                                                                                                                                                                                                                                                                                                                                                     Alignment Scores:
Pred. No.:
                                                                                                                                                                                                                            612
                                                                                                                                                                                                                                                                                                      US-10-133-133A-6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               RESULT 10
US-09-742-659-3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             182
                                                                                                                                                                                                                                                    TYPE: DNA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      22
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       42
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    139
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      62
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     661
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      259
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          439
                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Query Match:
DB:
                                                                                                                                                                                                                            LENGTH:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       QQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ő
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                đ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ద
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          D
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    οy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ογ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
```

```
3593
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             3773
                                                                                                                                                                                                                                                                                                                                                                                3473
                                                                                                                                                                                                                                                                                                                                                                                                                           3533
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    3653
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               113
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        133
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    153
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SerGlyGlyProLeuLeuCysProAlaGlyHisAlaValGlyIlePheArgAlaAlaVal 173
                                                                                                                                                                                                                                                                                                                                                                                                                                                  73
                                                                                                                                                                                                                                                                                                                                                                                                      53
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              93
                                                                                                                                                                                                                                                                                                                                                                     AspLysAspLeuValGlyTrpGlnAlaProGlnGlySerArgSerLeuThrProCysThr
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ::::::||| ::::::||| 3354 CGTAGGGCCAGGAGTAGTCTCCAAGGGGTGGAGG
                                                                                                                                                                                                                                                                                                                                                                                                                                                  54 AlaThrGlnThrPheLeuAlaThrSerIleAsnGlyValLeuTrpThrValTyrHisGly
                                                                                                                                                                                                                                                                                                                                                         ---LeuSerGlyAspThrAlaTyrAlaGlnGlnThrArgGlyGluGlnGlyThrGlnLys
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             74 AlaGlyThrArgThrIleAlaSerProLysGlyProValThrGlnMetTyrThrAsnVal
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                3654 GACCAAGACCTTGTGGGCTGGCCCGCTCTCAAGGTTCCCGGCTCATTGACACCCTGCACC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ArgGlyAspSerArgGlySerLeuLeuSerProArgProIleSerTyrLeuLysGlySer
                                                                                                                                                                                                       9646
168
10
17
       APPLICANT: Lau, Johnson Y.
APPLICANT: Lemon, Stanley M.
TITLE OF INVENTION: Chimeric HCV/GBV-B viruses
FILE REPRENCE: 1001116
CURRENT PELING DATE: 2000-12-21
PRIOR APPLICATION NUMBER: US 60/171,469
PRIOR PAPLICATION NUMBER: US 60/171,469
PRIOR PILING DATE: 1999-12-22
NUMBER OF SEQ ID NOS: 16
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 3
                                                                                                                                                                                                                               Conservative:
                                                                                                                                                                                                                                                                                       US-09-965-594-22 (1-197) x US-09-742-659-3 (1-9646)
                                                                                                                                                                                                                                                                                                           3 LysLysGlySerValValIleValGlyArgIleAsn-
                                                                                                                                                                                                                                           Mismatches:
                                                                                                                                                                                                         Length:
Matches:
                                                                                                                                                                                                                                                      Indels:
Gaps:
Wright-Minogue, Jacquelyn
Lau, Johnson Y.
                                                                                                                                                                                                      5.25e-88
848.50
87.25%
82.35%
83.51%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          3954 ATGAGATCCCCG 3965
                                                                                                                                                         ; ORGANISM: Hepatitis C virus US-09-742-659-3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             MetArgSerPro 197
                                                                                                                                                                                                                               Percent Similarity:
Best Local Similarity:
                                                                                                                                                                                                Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                154
                                                                                                                                                                                                                                                                                                                                                         15
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           56.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    134
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           194
 APPLICANT:
                                                                                                                                                   TYPE: DNA
                                                                                                                                                                                                                                                      Query Match:
DB:
                                                                                                                                                                                                            ..
9
                                                                                                                                                                                                         Pred.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                  a
                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                           ద
                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                        a
                                                                                                                                                                                                                                                                                                                                                                                                                                                õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                     a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              줨
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ם
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ద
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              õ
```

```
3894 IGCACCCGIGGAGTGGCTAAGGCGGTGGACTITAICCCTGTGGAGAACCTAGAGAGACAACC 3953
                                                                                                                                                                  94 AspLysAspLeuValGlyTrpGlnAlaProGlnGlySerArgSerLeuThrProCysThr 113
                                                         3654 GACCAAGACCTTGTGGGCTGGCCCGCTCCTCAAGGTTCCCGCTCATTGACACCCTGCACC
                                                                                                                 174 SerThrArgGlyValAlaLysAlaValAspPheIleProValGluSerLeuGluThrThr
                                                                                              114 CysGlySerSerAspLeuTyrLeuValThrArgHisAlaAspValIleProValArgArg
                                                                                                                                                                                                                                           154 SerGlyGlyProLeuLeuCysProAlaGlyHisAlaValGlyIlePheArgAlaAlaVal
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           TITLE OF INVENTION: FUNCTIONAL DNA CLONE FOR HEPATITIS C VIRUS (RCV) AND USES THEREOF
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: 0S/09/995,937
FILING DATE: 28-No. US20030028010A1-2001
CLASSIFICATION: <URNOWN>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CORRESPONDENCE ADDRESS:
ADDRESSEE: HOWELL & HAFERKAMP, L.C.
STREET: 7733 FORSYTH BLVD., SUITE 1400
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           REFERENCE/DOCKET NUMBER: 6029-4831
TELECOMMUNICATION INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/09/034,756
FILING DATE: 04-May-1998
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SEQUENCE DESCRIPTION: SEQ ID NO: 1:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          REGISTRATION NUMBER: 35,197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Sequence 1, Application US/09995937; Publication No. US20030028010A1; GENERAL INFORMATION: APPLICANT: RICE, CHARLES et al.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ATTORNEY/AGENT INFORMATION:
NAME: HOLLAND, DONALD R.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       LENGTH: 9646 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  relephone: 314-727-5188
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   TELEFAX: 314-727-6092
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 COMPUTER READABLE FORM:
                                                                                                                                                                                                                                                                                                                                                                                                               3954 ATGAGATCCCCG 3965
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               TOPOLOGY: linear MOLECULE TYPE: cDNA
                                                                                                                                                                                                                                                                                                                                                                                          194 MetArgSerPro 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                NUMBER OF SEQUENCES:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CITY: ST. LOUIS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              COUNTRY: USA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    HYPOTHETICAL: NO
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ZIP: 63105
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ANT I - SENSE :
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            US-09-995-937-1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   US-09-995-937-1
                                                                                                                                                                                                                                                                          a
                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                           8
                                                                                            ò
                                                                                                                             g
                                                                                                                                                                    δ
                                                                                                                                                                                                      g
                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       3354 CGTAGGGCCAGGAGATACTGCTTGGGCCAGCCGACGGAATGGTCTCCAAGGGGTGGAGG 3413
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     54 AlaThrGlnThrPheLeuAlaThrSerIleAsnGlyValLeuTrpThrValTyrHisGly 73
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ---LeuSerGlyAspThrAlaTyrAlaGlnGlnThrArgGlyGluGlnGlyThrGlnLys 33
                                      APPLICANT: RICE, CHARLES et al.

TITLE OF INVENTION: FUNCTIONAL DNA CLONE FOR HEPATITIS C
TITLE OF INVENTION: VIRUS (HCV) AND USES THEREOF
NUMBER OF SEQUENCES: 21
CORRESPONDENCE ADDRESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         3 LysLysGlySerValValIleValGlyArglleAsn.......
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      9646
168
10
17
9
                                                                                                                                                                                                                                                                            MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MN-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/238,076
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Length:
Matches:
Conservative:
Mismatches:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    US-09-965-594-22 (1-197) x US-09-238-076-1 (1-9646)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Indels:
                                                                                                                                              ADDRESSEE: HOWELL & HAFFRKAMP, L.C.
STREET: 7733 FORSYTH BLVD., SUITE 1400
CITY: ST. LOUIS
                                                                                                                                                                                                                                                                                                                                                                                                                             PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 09/034,756
Sequence 1, Application US/09238076
Patent No. US20020102540A1
GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                REGISTRATION NUMBER: 35,197
REGISTRATION NUMBER: 35,197
REFERENCE/COCKET UMBER: 607
TELECOMMUNICATION: 1NFORMATION:
TELECHONE: 314-727-5188
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   5.25e-88
848.50
87.25%
82.35%
83.51%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         TELEFAX: 314-727-6092
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ATTORNEY/AGENT INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     LENGTH: 9646 base pairs
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    TYPE: nucleic acid
STRANDEDNESS: double
                                                                                                                                                                                                                                                          COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       linear
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                             CLASSIFICATION:
                                                                                                                                                                                                                          USA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ANTI-SENSE: NO
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          MOLECULE TYPE:
HYPOTHETICAL:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                 FILING DATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Percent Similarity:
                                                                                                                                                                                                                                                                                                                                                                                            FILING DATE:
                                                                                                                                                                                                                                       ZIP: 63105
                                                                                                                                                                                                          ç
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Alignment Scores:
                                                                                                                                                                                                                        COUNTRY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                US-09-238-076-1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  15
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Query Match:
DB:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Best Local
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ..
⊗
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       õ
```

us-09-965-594-22.rnpb

```
Score:
                                                                                                                                                                                                                                                                               Pred
                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                           õ
                                                                                                                                                                                                                                                                                                                                                                                                                             a
                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                 qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         쉽
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               윱
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             셤
                                                                                                                                                                                                                                                                                                                                                 3773
                                                                                                                                                                                                                                                                                                                                                                   3533
                                                                                                                                                                                                                              3593
                                                                                                                                                                                                                                               3834 TCGGGGGGTCCGCTGTTGTGCCCCGCGGGACACGCCGTGGGCCTATTCAGGGCCGCGGGTG 3893
                                                                                                            3354 CGTAGGGGCCAGGAGATACTGCTTGGGCCAGCCGACGGATGGTCTCCAAGGGGTGGAGG 3413
                                                                                                                                                                                                                                                                                        113
                                                                                                                                                                                                                                                                                                                                                                                                           173
                                                                                                                                                                                                                                                                                                                               114 CysGlySerSerAspLeuTyrLeuValThrArgHisAlaAspValIleProValArgArg 133
                                                                                                                                                                                                                                                                                                                                                                                                                                                174 SerThrArgGlyValAlaLySAlaValAspPhcIleProValGluSerLeuGluThrThr 193
                                                                                                                                                                     53
                                                                                                                                                                                                           73
                                                                                                                               33
                                                                                                                                                                                AspLysAspLeuValGlyTrpGlnAlaProGlnGlySerArgSerLeuIhrProCysThr
                                                                                                                                                                                                                                                                                                                                         ---LeuSerGlyAspThrAlaTyrAlaGlnGlnThrArgGlyGluGlnGlyThrGlnLys
                                                                                                                                                                                                           54 AlaThrGlnThrPheLeuAlaThrSerIleAsnGlyValLeuTrpThrValTyrHisGly
                                                                                                                                                                     ThrSerHisThrGlyArgAspLysAsnGlnValGluGlyGluValGlnIleValSerThr
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      TITLE OF INVENTION: FUNCTIONAL DNA CLONE FOR HEPATITIS C VIRUS (HCY) AND USES THEREOF
9646
168
10
17
17
                    Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          NUMBER OF SEQUENCES: 21
CORRESPONDENCE ADDRESS:
ADDRESSEE: HOWELL & HAFERKAMP, L.C.
STREET: 7733 FORSYTH BIVD., SUITE 1400
CITY: ST. LOUIS
                                                                     JS-09-965-594-22 (1-197) x US-09-995-937-1 (1-9646)
                                                                                         LysLysGlySerValValIleValGlyArgIleAsn-
 Length:
Matches:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     COUNTRY: USA
ZIP: 63105
COMPUTER READABLE FORM:
WEDDUW TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
                                                    Gaps:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                et al
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Sequence 1, Application US/09917563
Publication No. US20030073080A1
GENERAL INFORMATION:
APPLICANT: RICE, CHARLES et
 5.25e-88
848.50
87.25%
82.35%
83.51%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  MetArgSerPro 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            STATE: MO
                     Percent Similarity:
Best Local Similarity:
Query Match:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              RESULT 13
US-09-917-563-1
                                                                                                                                                                                                                                                                   3594
                                                                                                                                                                                                                                                                                                                                                 3714
                                                                                         ٣
                                                                                                                                                                     34
                                                                                                                                                                                                                              3534
                                                                                                                                                                                                                                                 74
                                                                                                                                                                                                                                                                                       94
                                                                                                                                                                                                                                                                                                           3654
                                                                                                                                                                                                                                                                                                                                                                                                           154
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       194
  Pred. No.:
             Score:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         D
                                                                                                                                                                                                                                                                                                          g
                                                                                                        Q
                                                                                                                              ò
                                                                                                                                             g
                                                                                                                                                                    ò
                                                                                                                                                                                      d
                                                                                                                                                                                                          ö
                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                 qq
                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                   õ
                                                                                                                                                                                                                                                                                                                                                                                   a
                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                             a
                                                                                                                                                                                                                                                                                                                                                                                                                                               à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                å
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      õ
                                                                                         à
```

```
3354 CGTAGGGCCAGGAGTACTGCTTGGGCCAGCCGACGGAATGGTCTCCAAGGGGTGGAGG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ---LeuSerGlyAspThrAlaTyrAlaGlnGlnThrargGlyGluGlnGlyThrGlnLys
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        3414 TTGCTGGGGCCCATCACGGCGTACGCCCAGCAGAGGAGAGGCCTCCTAGGGTGTATAATC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         54 AlaThrGlnThrPheLeuAlaThrSerIleAsnGlyValLeuTrpThrValTyrHisGly
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CysGlySerSerAspLeuTyrLeuValThrArgHisAlaAspValIleProValArgArg
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        34 ThrSerHisThrGlyArgAspLysAsnGlnValGluGlyGluValGlnIleValSerThr
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              AspLysAspLeuValGlyTrpGlnAlaProGlnGlySerArgSerLeuThrProCysThr
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  3714 TGCGGCTCCTCGGACCTTTACCTGGTCACGAGGCACGCCGATGTCATTCCCGTGCGCGG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  9646
168
10
17
9
<u>.</u>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Length:
Matches:
Conservative:
Mismatches:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      3 LysLysGlySerValValIleValGlyArgIleAsn--
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        US-09-965-594-22 (1-197) x US-09-917-563-1 (1-9646)
                                                                                                                                                                                  6029-4831
                            APPLICATION NUMBER: US/09/917,563
FILING DATE: 27-Jul-2001
CLASSIFICATION: <URNOWN>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        indels:
                                                                             PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/238,076
FILING DATE: 26-JAN-1999
                                                                                                                                                                                                                                                                                                                                                                                                     NO: 1:
                                                                                                                            ATTORNEY/AGENT INFORMATION:
NAME: HOLLAND, DONALD N:
REGISTRATION UNGBER: 35,197
REFERENCE/DOCKET NUMBER: 60:
                                                                                                                                                                                           TELECOMMUNICATION INFORMATION:
TELEPHONE: 314-727-5188
TELEPHONE: 314-727-5092
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 9646 Dasse pairs
TYPE: nucleic acid
STRANDEDNESS: double
                                                                                                                                                                                                                                                                                                                                                                                  ANTI-SENSE: NO
SEQUENCE DESCRIPTION: SEQ ID
              CURRENT APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                   5.25e-88
848.50
87.25%
82.35%
83.51%
                                                                                                                                                                                                                                                                                                                                  TOPOLOGY: linear MOLECULE TYPE: CDNA
                                                                                                                                                                                                                                                                                                                                                                    HYPOTHETICAL: NO
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Best Local Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Percent Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                     Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                  US-09-917-563-1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              46
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  114
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        3834
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        12
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       154
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Query Match:
DB:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ..
∑
```

3473

33

53

3533

3593

93

73

3653

3713

113

3773

133

3833

```
3774 CGAGGTGATAGCAGGGTAGCCTGCTTTCGCCCCGGCCCATTCCTACTTGAAAGGCTCC 3833
                                                                                                                                                                                                                                                                                                                                                                                                                                    3533 ACCAGCCIGACIGGCCGGGACAAAACCAAGIGGAGGGIGAGGICCAGAICGIGICAACI
                                                                                             3534 GCTACCCAAACCTTCCTGGCAACGTGCATCAATGGGGTATGCTGGACTGTCTACCACGGG 3593
                                                                                                                                                                                                                                ArgGlyAspSerArgGlySerLeuLeuSerProArgProlleSerTyrLeuLysGlySer 153
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               174 SerThrArgGlyValAlaLysAlaValAspPheIleProValGluSerLeuGluThrThr 193
                                                                                                                                   63
                                                                                                                                                                                                           94 AspLysAspLeuValGlyTrpGlnAlaProGlnGlySerArgSerLeuThrProCysThr
                                                       AlaThrGlnThrPheLeuAlaThrSerIleAsnGlyValLeuTrpThrValTyrHisGly
                                                                                                                                 74 AlaGlyThrArgThrIleAlaSerProLysGlyProValThrGlnMetTyrThrAsnVal
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  IITLE OF INVENTION: FUNCTIONAL DNA CLONE FOR HEPAIITIS C VIRUS (HCV) AND USES THEREOF
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         COMPUTER: IBM PC compatible
COMPUTER: IBM PC compatible
SOFTWARE: Patentin Release #1.0, Version #1.30
SOFTWENT PATENTIN Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/995,937
FILING DATE: 28 No. US20030028010al-2001
CLASSITGATION ATA:
APPLICATION NUMBER: US/09/034,756
FILING DATE: 04-May-1998
ATTORNEY/ACENT INFORMATION:
NAME: HOLLAND, DONALD R.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CORRESPONDENCE ADDRESS:
ADDRESSE: HOWELL & HAFERRAMP, L.C.
STREET: 7733 FORSYTH BLVD., SUITE 1400
CITY: ST. LOUIS
STATE: MO
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              REGISTRATION NUMBER: 35,197
REFERENCE/DOCKET NUMBER: 6029-4831
TELECOMMUNICATION INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      : Sequence 5, Application US/09995937; Publication No. US20030028010A1
GENEAL INFORMATION: GENERAL INPORMATION: APPLICANT: RICE, CHARLES et al.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   COUNTRY: USA
ZIP: 63105
COMPUTER READABLE FORM:
MEDIUM TYPE: FLOPPY disk
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SEQUENCE CHARACTERISTICS:
LENGTH: 12980 base pairs
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    TELEPHONE: 314-727-5188 TELEFAX: 314-727-6092 INFORMATION FOR SEQ ID NO: 5:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1111111111
3954 ATGAGATCCCCG 3965
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         194 MetArgSerPro 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SEQUENCES:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          NUMBER OF
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  RESULT 15
US-09-995-937-5
                                                                                                                                                                                                                                                                                                                                                              134
                                                       54
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                         qq
                                                                                                                               ŏ
                                                                                                                                                             g
                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                          셤
                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                    q
                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                         셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           à
                                                         9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    3414 TTGCTGGCGCCCATCACGGCGTACGCCCAGCAGACGAGGCCTCCTAGGGTGTATAATC 3473
                 3894 TGCACCCGTGGAGTGGCTAAAGCGGTGGACTTTATCCCTGTGGAGAACCTAGAGAACA 3953
174 SerThrArgGlyValAlaLysAlaValAspPhelleProValGluSerLeuGluThrThr 193
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ---LeuSerGlyAspThrAlaTyrAlaGlnGlnThrArgGlyGluGlnGlyThrGlnLys 33
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 34 ThrSerHisThrGlyArgAspLysAsnGlnValGluGlyGluValGlnIleValSerThr 53
                                                                                                                                                                                   Sequence 5, Application US/09238076
Patent No. US20020102540A1
GENERAL INFORMATION
TITLE OF INVENTION: FUNCTIONAL DNA CLONE FOR HEFATITIS C
TITLE OF INVENTION: VIRUS (HCV) AND USES THEREOF
NUMBER OF SEQUENCES: 21
CORRESPONDENCE ADDRESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                   MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PATENTIN Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/238,076
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          LysLysGlySerValVallleValGlyArglleAsn---
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Length:
Matches:
Conservative:
Mismatches:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      US-09-965-594-22 (1-197) x US-09-238-076-5 (1-12980)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Indels:
                                                                                                                                                                                                                                                                                                                                      ADDRESSEE: HOWELL & HAFERKAMP, L.C.
STREET: 7733 FORSYTH BLVD., SUITE 1400
CITY: ST. LOUIS
STATE: MO
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  NAME: HOLLAND, DONALD R.
REGISTRATION NUMBER: 35,197
REFERENCE/POCKET NUMBER: 6029-4831
TELEPHONE: 314-727-5188
TELEPHONE: 314-727-5188
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 09/034,756
FILING DATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        TELEFAX: 314-727-6092
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   7.72e-88
848.50
87.25%
82.35%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              : 12980 base pairs
nucleic acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ATTORNEY/AGENT INFORMATION
                                                                                               3954 ATGAGATCCCCG 3965
                                                                           MetArqSerPro 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   double
                                                                                                                                                                                                                                                                                                                                                                                                                                                       COMPUTER READABLE FORM:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     linear
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   TOPOLOGY.

MOLECULE TYPE: CD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Best Local Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               STRANDEDNESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ANTI-SENSE: NO
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Percent Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             FILING DATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                    63105
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Alignment Scores
                                                                                                                                                                                                                                                                                                                                                                                                                  COUNTRY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             US-09-238-076-5
                                                                                                                                                     RESULT 14
US-09-238-076-5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                LENGTH:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      15
                                                                           194
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Query Match:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ŏ
                                                                                                               g
                                   g
                                                                           ò
```

```
3354 CGTAGGGCCAGGAGATACTGCTTGGGCCAGGCGGAATGGTCTCCAAGGGGTGGAGG 3413
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       3473
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SerGlyGlyProLeuLeuCysProAlaGlyHisAlaValGlyIlePheArgAlaAlaVal 173
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     74 AlaGlyThrArgThrIleAlaSerProLysGlyProValThrGlnMetTyrThrAsnVal 93
                                                                                                                                                                                                                                                                                                                                                                                                                                        4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              33
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            AlaThrGlnThrPheLeuAlaThrSerIleAsnGlyValLeuTrpThrValTyrHisGly 73
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            LysLysGlySerValVallleValGlyArgIleAsn--------------
                                                                                                                                                                                                                             12980
168
10
17
9
                                                                                                                                                                                                                             Length:
Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                          US-09-965-594-22 (1-197) x US-09-995-937-5 (1-12980)
                                                                                                                                                                                                                                                                                                                                                 Gaps:
                                                                                                                                      ..
د.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           completed: August 31, 2003, 04:55:02
                                                                                                                                         ..
O
                                                                                                                                         ΙD
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: CDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
SEQUENCE DESCRIPTION: SEQ I)
                                                                                                                                                                                                                                7.72e-88
848.50
87.25%
82.35%
83.51%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ATGAGATCCCCG 3965
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            MetArgSerPro 197
                                                                                                                                                                                                                                                          Score:
Percent Similarity:
Best Local Similarity:
Query Match:
DB:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Search completed: Augus
Job time : 189.482 secs
                                                                                                                                                                                                          Alignment Scores:
Pred. No.:
       TYE

STH

TOG

MOLECULE

HYPOTHEN

ANTI-SER

SEQUENCE

13 SEQUENCE

16 SEQUENCE

17 SECUENCE

18 SEQUENCE

18
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   3594
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            3654
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          3834
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 $6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          3954
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            154
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              194
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              24
                                                                                                                                                                                                                                                                                                                                                                                                                                        ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       á
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Q
D
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         οŽ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                සි සි
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  S G
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          S a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            õ
```

,\*

GenCore version 5.1.6 Copyright (c) 1993 - 2003 Compugen Ltd.

```
DEFINITION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ACCESSION
VERSION
KEYWORDS
SOURCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ORGANISM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      RESULT 1
CB950999
                                                                                                                                                                                                                                                                                                                                                     (without alignments)
2506.388 Million cell updates/scc
                                                                              US-09-965-594-22
1016
1 MKKKGSVVIVGRINLSGDTA......VAKAVDFIPVESLETIMRSP 197
                          August 30, 2003, 19:20:43; Search time 1910.31 Seconds
                                                                                                                                                                                                                                          45562784
OM protein - nucleic search, using frame_plus_p2n model
                                                                                                                                                                                                              22781392 seqs, 12152238056 residues
                                                                                                                                                                                                                                         Total number of hits satisfying chosen parameters:
                                                                                                                              BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0
                                                                                                                                                                                                                                                                                                       Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            em_gss_pro:
em_gss_rod:
em_gss_phg:
em_gss_vrl:
gb_gssl:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        em_gss_vrt:*
em_gss_fun:*
em_gss_mam:*
                                                                                                                                                                                                                                                                  Minimum DB seq length: 0
Maximum DB seq length: 2000000000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  em_gss_hum:*
em_gss_inv:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              em_esthum: *
em_estin: *
em_estin: *
em_estin: *
em_estp: *
em_estp: *
em_btc: *
gb_est: *
gb_est: *
gb_htc: *
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        em_estfun:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             em_gss_pln:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  em_gss_mus:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    em_estom:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            gb_est4:*
gb_est5:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 em_estba:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  gb_est3:*
                                                                               Title:
Perfect score:
                                                                                                                                   Scoring table:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Database :
                                                                                                        Sequence:
                                                                                                                                                                                                               Searched:
                           Run on:
```

# 29: qb\_gss2:\*

		ď			SUMMARIES		
Resul	t Score	ery	Length	<b>B</b> O	ID	ipti	
	8	10.3	1031	14	CB95099	B950999	
	100.		40	13.	B0925101	0610	AGENCOUR
o	5		64	5 6	BX238988	100	1 -
	97.		~	10	BG089727	8972	9
	σ		4	12	BM915803	1580	AGENCOUR
	95.		70	10	BF863244	-	963042C
Ü	8 95.5		~	29	CC406704	670	4 PUHKD12TB
,	95.		æ	53	CC406705	0	PUHKD121
o O	95.	٠	8	10	BF304699	Ġ.	60188825
٦,	ָ ק	٠	70	53	B2342381	8	1c83b11
- 1	4.		9	53	CNSO60HN	~	T3 end
٦.	4.		4	15	BM467279	_	AGENCOUR
٠,	ית	•	~	12	BM402566	ā	SLA005FJ
٦,	ייס	٠	20	78	AQ538021	N	RPCI-11
U	6	•	-	13	BU541777	'n	AGENCOU
-	92.	•	93	13	BQ894657	in	AGENCOUR
٦.	92.	٠	∾.	53	N	<u></u>	pacs2-1(
~1 Ĉ	י ע	•	~ (	20	BE055938	m,	945020DC
<b>N</b> (	2	٠	8		0	'n	AL571605
4 C			N H	ם זר	AC620249		HS_5182
יוני טנ			7 0		D6420660	0 5	50245200
10			"		DEC00113		POSTOTO
4 C	. 0	٠	ם ע		#/10000E	. 7	AGENCOO!
• (	no	٠	<b>ص</b> ر		D1459445		WOLLC DA
4 (	no	•	າປ				6032004
	10	•	١ ٧		D10202013	4 -	TACTACO
10	10	٠	<b>)</b> (		044243	• 0	0001000
יה ט כ	no	•	> -		0711169	οú	DA4 34 20
יר	,	٠	٠,		B1311100	ŏ	003002/
<b>~</b>	900		a a		81329116	_	60298028
n (	96.	٠	<b>`</b> '		BZ55232/	~	pacs1-6(
<b>v</b> ) (	90.		Н.		BU299264	õ	60360843
m	90.		Г		NS0917	Ö	Single
m	90.		ч		43	~	60332307
č O	96		C	13	039692	Ñ	60340500
3	90.		7		36575	16575	6035847
m	90.		C	13	9345	345	AGENCOITET
m	.06		10	13	BU956626	5662	FNCOT
4	90.		**	1	3054	8054	Mus musco
4	6		0	12	800	000	T
0	6		9	12	2531	5531	O E V.
4	σ		o	10	3275	2000	3000
4	ν σ		١,٠	9 5	7105	200	0000
•	n (	٠	1	7	2010	5010	212
			L	,			

EST 29-APR-2003

CB950999 1031 bp mRNA linear EST AGENCOURT\_13445496 NIH\_MGC\_177 Mus musculus cDNA clone IMAGE:30316162 5', mRNA sequence.

CB950999.1 GI:30205777

CB950999

ALIGNMENTS

EST.

Mus musculus (house mouse)

Mus musculus

Eukaryota, Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutharia; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 (bases 1 to 1031)

REFERENCE

N

```
175 rArgGlyValAlaLysAlaValAspPheIleProValGluSerLeuGluThrThrMetAr 195
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     /tissue_type="sympathetic trunk"
/dev_stage="adult, 16 yr"
/lab_host="DH10B"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Conservative:
Mismatches:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             211 t
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Matches:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1. .1199
forganism".Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="INAGE:6192708"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Gaps:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          US-09-965-594-22 (1-197) x BQB92487 (1-1199)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             343 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Technólogies."
362 c 34
                                                                                                                                                                                                                                                                                                                                                                                               BQ892487.1 GI:22284501
EST.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Homo sapiens (human)
Homo sapiens
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            /sex-'male'
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         101.00
34.74%
23.47%
9.94%
                                                                                                                   195 gSerPro 197
                                                                                                                                                     772 C-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Percent Similarity:
Best Local Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                255
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   source
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          BASE COUNT
ORIGIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Pred. No.:
                                                                                                                                                                                                                                                                                                                   DEFINITION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ORGANISM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              AUTHORS
TITLE
JOURNAL
COMMENT
                                                                                                                                                                                                                                                                                                                                                                                                        VERSION
KEYWORDS
SOURCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      REFERENCE
                                                                                                                                                                                                                                 RESULT 2
BQ892487
                                                                                                                                                                                                                                                                                                                                                                                ACCESSION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     FEATURES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Score:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                      õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      /Libe_invalle_invalle_invalle_invalle_invalle_invalle_invalle_invalle_invalle_invalle_invalle_invalle_invalle_invalle_invalle_invalle_invalle_invalle_invalle_invalle_invalle_invalle_invalle_invalle_invalle_invalle_invalle_invalle_invalle_invalle_invalle_invalle_invalle_invalle_invalle_invalle_invalle_invalle_invalle_invalle_invalle_invalle_invalle_invalle_invalle_invalle_invalle_invalle_invalle_invalle_invalle_invalle_invalle_invalle_invalle_invalle_invalle_invalle_invalle_invalle_invalle_invalle_invalle_invalle_invalle_invalle_invalle_invalle_invalle_invalle_invalle_invalle_invalle_invalle_invalle_invalle_invalle_invalle_invalle_invalle_invalle_invalle_invalle_invalle_invalle_invalle_invalle_invalle_invalle_invalle_invalle_invalle_invalle_invalle_invalle_invalle_invalle_invalle_invalle_invalle_invalle_invalle_invalle_invalle_invalle_invalle_invalle_invalle_invalle_invalle_invalle_invalle_invalle_invalle_invalle_invalle_invalle_invalle_invalle_invalle_invalle_invalle_invalle_invalle_invalle_invalle_invalle_invalle_invalle_invalle_invalle_invalle_invalle_invalle_invalle_invalle_invalle_invalle_invalle_invalle_invalle_invalle_invalle_invalle_invalle_invalle_invalle_invalle_invalle_invalle_invalle_invalle_invalle_invalle_invalle_invalle_invalle_invalle_invalle_invalle_invalle_invalle_invalle_invalle_invalle_invalle_invalle_invalle_invalle_invalle_invalle_invalle_invalle_invalle_invalle_invalle_invalle_invalle_invalle_invalle_invalle_invalle_invalle_invalle_invalle_invalle_invalle_invalle_invalle_invalle_invalle_invalle_invalle_invalle_invalle_invalle_invalle_invalle_invalle_invalle_invalle_invalle_invalle_invalle_invalle_invalle_invalle_invalle_invalle_invalle_invalle_invalle_invalle_invalle_invalle_invalle_invalle_invalle_invalle_invalle_invalle_invalle_invalle_invalle_invalle_invalle_invalle_invalle_invalle_invalle_invalle_invalle_invalle_invalle_invalle_invalle_invalle_invalle_invalle_invalle_invalle_invalle_invalle_invalle_invalle_invalle_invalle_invalle_invalle_invalle_invalle_invalle_invalle_invalle_invalle_
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    104 Gln-------GlySerArgSerLeuThrProCysThrCysGlySerSerAsp 118
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 119 LeuTyrLeuValThrArgHisAlaAsp~ValIleProValArgArgArgGlyAspSerAr 138
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     651
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            138 gGlySerLeuLeuSerProArgProlleSerTyrLeuLysGlySerSerGly----- 155
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                -GlyProLeuLeuCysProAlaGlyHisAlaValGlyIlePheArgAlaAlaValSerTh 175
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  395 ATTCAAGGCTATCCCCAAAACAAGAGAGTACATCGGCAGCTTTTCCTT---CACTCATTT 451
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          84 GlyProValThrGlnMetTyrThrAsnValAspLysAspLeuValGlyTrpGlnAlaPro 103
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            44 ValGluGlyGluValGlnIleValSerThrAlaThrGlnThrPheLeuAlaThrSerIle 63
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Email: cgapbs-remail.nih.gov
Tissue Procurement: Dr. Michael Brownstein
cDNA Library Preparation: Michael Brownstein Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov.b column: 11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         64 AsnGlyValLeuTrpThrValTyrHisGlyAlaGlyThrArgThrIleAlaSerProLys
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            604 ------ACCACCCATGGGGTTGTTGTTCCCGGCTTCTCCCGTGGGCAATTACA
NIH-MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1031
51
16
62
35
8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Length:
Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                1. .1031
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="lMAGE:30316162"
                                                                                                                                                                                                                                                                                                                                     Plate: NDCM107 row: b column: 11
High quality sequence stop: 333.
Location/Qualifiers
                                                                                        Contact: Robert Strausberg, Ph.D.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  US-09-965-594-22 (1-197) x CB950999 (1-1031)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               3.53
104.50
41.10%
31.29%
10.29%
                                                            Unpublished
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Best Local Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 235
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Percent Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Query Match:
DB:
                                                                                                                                                                                                                                                                                                                                                                                                                                          source
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           BASE COUNT
ORIGIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Pred. No.:
AUTHORS
TITLE
JOURNAL
COMMENT
                                                                                                                                                                                                                                                                                                                                                                                                           FEATURES
```

g

ò g ò g ð

g à g ö

셤 õ

```
//Clone_11b="Lupski_sympathetic_trunk"
//Clone_11b="Lupski_sympathetic_trunk"
//note="Vector: pcMv-SPORT6 (Life Technologies); Site_1:
Not1; Site_2: Sali; cDNA made by oligo-dT priming.
Directionally cloned using the following adaptors:
5'-TCGACCCAGCGTCCG-3' and
5'-TCGACCCAGCGTCCG-3' and
5'-GACTAGTTCTAGATCGCGGCGCCCCT(15)-3'. Size selected >
1'Ab for average insert length 1.9 kb. This is a primary
library, non-amplified. Library constructed by Life
Technologies and donated by J. Lupski, M.D./Ph.D. (Baylor
College of Medicine); available through Life
                                                                                                                                                                                                                            1199 bp mRNA linear EST 16-AGG-2002 mRNA linear EST 16-AGG-2002 IMAGE:6192708 5', mRNA sequence.
-----AAGGCA 369
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           16 SerGlyAspThrAlaTyrAlaGlnGlnThrArgGlyGluGlnGlyThrGlnLysThrSer 35
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Eukaryota: Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1199)
NIH-MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC) Unpublished
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Dr. James R. Lupski
CDNA ilbrary Preparation: Life Techhologies, Inc.
CDNA ilbrary Preparation: Life Techhologies, Inc.
CDNA ilbrary Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bloscience Corporation
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LiAMISSS row: c column: 13
High quality sequence start: 57
High quality sequence stop: 394.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          þe
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  28 others
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    GCAGGAGAGAACCTTACCCCAAACAG--------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1199
50
24
80
59
```

us-09-965-594-22.rst

/db_xref="taxon:10090" /clone="IMAGE:6335718" /lab_host="DH10B (phage-resistant)" /clone_lib="NIH_MGC_130" /clone_lib="NIH_MGC_130" /note="Organ: otocysts; Vector: pCMV-SPORT6.1.ccdb; Site=1: EccRV; Site=2: Not1; Cloned unidirectionally. Primer: Oligo dT. Average insert size 1.95 kb. Constructed by ResGen, Invitrogen Corp. Note: this is a NIH_MGC Library."  BASE COUNT 297 a 521 c 237 g 345 t 3 others ORIGIN	Alignment Scores: 13.2 Length: 1403 Pred. No.: 100.50 Matches: 56 Score: 100.50 Matches: 56 Best Local Similarity: 27.86% Mismatches: 74 Query Match: 13 Gaps: 56 DB: 13 Gaps: 9 US-09-965-594-22 (1-197) x BQ926101 (1-1403)	4 Description of the state of t	1264 ACAGAGGGAAA	Qy         103 rodinGlySerAegSerLeuthrFroCyaflySerSerAapLeuTyrLeuValT 123           1147 AcchaGagaGCaCCTGACGCCCTGTGGGCNTGTCGATAATAACAAATGTG 1091           123 hrargHisalaAspVall1eProValargArgGlyAsp	PD 922 CGCAGGAGAGGCGGCGTGTTTCGGCGGGTGAGCGCGCGGTG 875 RESULT 4 BX238988/C LOCUS BX238988 6 644 bp DNA linear GSS 29-JAN-2003 DEFINITION Danio rerio genomic clone DKEY-283L13, genomic survey sequence. ACCESSION BX238988 1 GI:28161322 KEYMORDS GSS. SOURCE Danio rerio (zebrafish) ORGANISM Danio rerio Eukaryota: Metazoa; Chordata; Craniata; Vertebrata; Euteleostom1: Actinopteryg11; Neopteryg11; Teleoste1; Ostariophysi;
36 HisThrGlyArgAspLysAsnGlnValGluGlyGluValGlnIleValSerThr 53	Db	Db   637 ATTGGCCACTCCTCTATAAAACGCCAACGTCGTTCCATGGGGGGGTGGGT	Oy 157 oLeuLeuCysProAlaGlyHisAlaValGlyIlePheArgAlaAlaValSerThrArgGl 177  Db 817 CCTAACATCTCCTACCCGGGGGGGGGGGGGGGGGGGGGG	HUGUSTON AGENCOURT_8752655 NIH_MGC_130 Mus musculus CDNA clone IMAGE:6335718  SCHENITION AGENCOURT_8752655 NIH_MGC_130 Mus musculus CDNA clone IMAGE:6335718  SCHENITION AGENCOURT_8752655 NIH_MGC_130 Mus musculus CDNA clone IMAGE:6335718  ACCESSION BQ926101.1 GI:22341132  KEYWORDS EST.  KEYWORDS EST.  BUSCALLUS (house mouse)  ORGANISM Mus musculus (house mouse)  ORGANISM Mus musculus (house mouse)  AUTHORISM HIW MC http://mgc.nci.nih.gov/.  AUTHORS INH MGC http://mgc.nci.nih.gov/.  TITLE Unpublished Contact: Robert Strausberg, Ph.D.  COMMENT Contact: Robert Strausberg, Ph.D.  Tissue Procurement: Mark Maconochie, Ph.D. and Nancy L. Freeman,	ion (LLA

```
Tumor Gene Index
Unpublished
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Best Local Similarity:
Query Match:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Percent Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  36
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         56
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          578
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             74
                                                                                                                                                                                                                                                                                                                                                                                                                                                                             source
                                 VERSION
KEYWORDS
SOURCE
ORGANISM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ..
Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        BASE COUNT
                                                                                                                                                         AUTHORS
TITLE
                                                                                                                                                                                                         JOURNAL
COMMENT
                ACCESSION
                                                                                                                                         REFERENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                          FEATURES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ORIGIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Pred.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         BG089727 62-JAN-2001 mRNA linear, EST 26-JAN-2001 mab90e06.xl NCI_CGAP_Sp2 Mus musculus cDNA clone IMAGE:3977578 3' similar to SW:GRAD_MOUSE Pl1033 GRANZYME D PRECURSOR;, mRNA
              Humphray, S.J., Huckle, E. and Durham, J.L.

Birect Submitssion

Submitted (27-JAN-2003) The Sanger Institute, Wellcome Trust Genome Campus, Hinxton, Cambridgeshire, CBIO 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk Unpublished

This sequence was generated from the T7 end of BAC 283L13. 283L13 is part of the banjokey BAC Library created by R. Plasterk and N.V. Keygene. Further details:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        146
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             108
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  108 rLeuThrProCysThrCysGlySerSerAspLeuTyrLeuValThrArgHisAlaAspVa 128
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              147 -IleSerTyrLeuLysGlySerSerGlyGlyProLeuLeu---CysProAlaGlyHisAl 165
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              316 TCTTGGAAAGGAC-----TGGAGATCCAGACCGTTGGGGACAAAGGCGGAGGAGTC 266
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       GAGACTGCATGGACCACCGGCGTGCCGACGCTGGACACCATCCTGGCAGAGTGGGAGCTC 377
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            376 rcgaggcccrgrccgccgrcaccagrgaccagcrgcrgrcargrcggrggrcagcag 317
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  GACCATGGCTATTACCATCACTGGAGGGGGGGGGCAGCTCGGAATCCGGACGCGCCATCTC 497
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    196 CTTGAGCGGTGCAGCGGACCGGAAGCCTTCCCTGGCCCGGGTGCAGAGTGCCATTAGGGCC 437
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          92
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   57
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      58 PheLeuAlaThrSerlleAsnGlyValLeu-TrpThrValTyrHisGlyAlaGlyThrAr 77
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                18 AspThrAlaTyrAlaGlnGlnThrArgGlyGluGlnGlyThrGlnLysThrSerHisThr 37
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        77 gThrIleAlaSerProLysGlyProValThrGlnMetTyr------ThrAs
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 nValAspLysAspLeuValGlyTrpGlnAla---ProGlnGlySerArg-----Se
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   265 GICACCACCATGCCCTITAAAAAAGGAGAAAITATTTGTGATTACCATGGGGAGGAGTTG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          128 IIleProValArgArgGlyAspSerArgGlySerLeuLeuSerProArgPro----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        205 TCGCACGAGGAGGCTTAGGGAGACGGGAGGCCCTCCCCCGTGTCTCCTCCTCCTACATG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     GlyArgAspLysAsnGlnValGluGlyGluValGInIleValSerThrAlaThrGlnThr
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   165 aValGlyIlePheArgAlaAlaValSerThrArgGlyVal 178
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               TGCCCTGTCACCCGGGCATGGAAACCTTCGGAAGGCTTA 49
                                                                                                                                                                                                                                                                                                                                                                                                        644
48
25
84
17
                                                                                                                                                                          http://www.sanger.ac.uk/Projects/D_rerio/
Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                       Length:
Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                          /organism="Danio rerio"
//mol_type="genomic DNA"
//db_xref="taxon:7955"
/clone="DKEY-283L13"
/fissuc_type="Testis"
/note="vector pindigoBAC-536"
a 212 c 176 q 127 t
Cypriniformes; Cyprinidae; Danio.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Gaps:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               US-09-965-594-22 (1-197) x BX238988 (1-644)
                                                                                                                                                                                                                                                                                                                                                                                                       8.45
98.00
41.958
27.598
9.658
                                                                                                                                                                                                           .644
                                                                                                                                                                                                                                                                                                                                                                                                                                          Percent Similarity:
Best Local Similarity:
                                                                                                                                                                                                                                                                                                                                                                                        Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             LOCUS
DEFINITION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Query Match:
DB:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         RESULT 5
BG089727/c
                                                                                                                                                                                                                                                                                                                                        SASE COUNT
                                                                                                                                                                                                                                                                                                                                                                                                              ..
Q
                                      AUTHORS
                                                                        JOURNAL
                                                                                                                                                                                              FEATURES
                                                                                                                            COMMENT
                                                                                                                                                                                                                                                                                                                                                        ORIGIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
```

```
//mol_type="mrnA"
//mol_type="mrnA"
/db_xref="taxon:10090"
/clone="landE:397578"
/tissue_type="nc cells (flow-sorted)"
/lab_host="DH10B (Tl-resistant)"
/lab_host="DH10B (Tl-resistant)"
/clone=lib="NCLCGAP_$92"
/note="Organ: spleen; yector: pcNV-SpORT6 (Life
Technologies); mRNA made from flow-sorted NK cells, cDNA
made by oligo-dT priming. Directionally cloned. Average
insert size 1.5 kb. Primary library, non-amplified. cDNA
Library Preparation: David B. Kizman, Ph.D."
illorary Preparation: David B. Kizman, Ph.D."
                                                                                                                                                                                                                                                                                                                                                            Email: cgapbs-remain.nin.gov
Tissue Procurement: David Segal Ph.D., Herbert Morse M.D.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Mashington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         413 ------GGGTCAAGGTCCATCAATGACACTAAAGCATCTGCCCGCCTGCGAGGGTT 363
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 AAAACATTTCCCATTCCAGATATAAATGCT-----ACTATCTTCTCAGGTGACATC 528
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               527 ATGCTGTTAAAGCTGGAGAGTAAGGCCAAGAGAACTAAAGCTGTGAGACCCCTCAAGTTG 468
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            AlaProGlnGlySerArgSerLeu----ThrProCysThrCysGlySerSerAspLeu 119
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ProLysGlyProValThrGlnMetTyrThrAsnValAspLysAspLeuValGlyTrpGln 101
                                                                                                                                         Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalla; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. I Chases 1 to 629)
NCI-GGAP Http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (GGAP),
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          GlnThrPheLeuAlaThrSerIleAsnGlyValLeuTrpThrValTyrHisGly~~---
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ---- AlaGlyThrArgThrIleAlaSer
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      629
23
23
59
121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Length:
Matches:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Seq primer: -400P from Gibco
High quality sequence stop: 422.
Location/Qualiflers
    1. .629
/organism="Mus musculus"
                                                                                                                                                                                                                                                                                                                                       Contact: Robert Strausberg, Ph.D.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    US-09-965-594-22 (1-197) x BG089727 (1-629)
                                                                                      Mus musculus (house mouse)
Mus musculus
BG089727
BG089727.1 GI:12572290
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      9.19
97.50
39.23%
26.52%
9.60%
```

ch: 9.45% Indels: 57  -594-22 (1-197) x BM915803 (1-1146)  22 AlaGlnGlnThrArgGlyGluGlnGlyThrGlnLysThrSer-HisThrGlyArgAspLy 41	948 AGCGGAGGCGCCGCGCGCGCGCGCGCGGGGGGCGCGCGGGCGGGGCGC	127 pvallleProvalargArgArgArgArgArgArgClySerLeuLeuSerProArgProll 147	BE863244  96304202.x1 C. reinhardtii CC-1690, Stress condition I, normalized Lambda Zap II Chlamydomonas reinhardtii cDNA, mRNA sequence. BF863244  BF863244.1 GI:12253388  EST. Chlamydomonas reinhardtii Lefebvre, Viridiplantee: Chlorophyta; Chlorophyceae; Volvocales; Chlamydomonadaceae: Chlamydomonas. I Chases I to 701) Grossman, A., Davies, J., Federspiel, N., Harris, E., Hauser, C., Lefebvre, P., McDermott, J. P., Shrager, J., Sliflow, C. and Stern, D. Analyses of the Chlamydomonas reinhardtii Genome: A Model	Vascular Plants; project phase 3 vene function and regulation Unpublished Contact: Charles Hauser Contact: Charles Hauser Contact: Charles Hauser Durkam, NC 27708-1000 Tel: 919 613 8159 Tel: 919 613 8159 Email: chauser@duke.edu. Location/Qualifiers 1701
Query Match: DB: US-09-965-594 QY 22 QY 41 QY 41 Db 1038 QY 61 Db 1002	ବ ୪ ବ ୪ ବ	6 6 6 6 6 6 6	RESULT 7 BF863244 LOCUS DEPINITION ACCESSION VERSION KEYWORDS SOURCE ORGANISM AUTHORS TITLE	JOURNAL COMMENT COMMENT FEATURES SOURCE
120 TyrLeuvalThrargHisalaaspValIleProvalArgArgArg	4 .	_	http://image.llnl.gov Plate: LLCM2007 row: i column: 01 Plate: LLCM2007 row: i column: 01 High quality sequence start: 6 High quality sequence stop: 256. Location/Qualifiers 1l146 /organism="Homo sapiens" //moltype="mmna" /db xref="taxon:9606" /clone="IMAGE:5482056" /clone="IMAGE:5482056" /lab bost="ThAGE:5482056" /lab bost="ThAGE:5482056" /lab bost="ThAGE:5482056" /lab bost="Sin: Vector: porgs: Site_1: XhoI; Site_2: EcoRI; cDNA made by oligo-dT priming: Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGACAGGGG(). Library constructed by Ling Hong in the laboratory of Garal M Dubis (Iniversity of California	
	RESULT 6 BM915803/c LOCUS DEFINITION ACCESSION KEYNORDS COURTER	SOURCE ORGANISM AUTHORS TITLE JOURNAL COMMENT	FEATURES	BASE COUNT 1 ORIGIN Alignment Scores: Score: Percent Similarit Best Local Similarit

us-09-965-594-22.rst

BASE COUNT

ORIGIN

.. Q

ò 셤 ò g ò qq ò QQ õ g õ 셤 ò

```
GlnGlySerArgSerLeuThrProCysThrCysGlySerSerAspLeu------ 119
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                334 CGRGGGGTTACTTTGGAACCCACACCAGGTCGCCATGTCAAGGTGGCAAATGGAGACCCA 275
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     158 LeuLeuCysProAlaGlyHisAlaValGlyIlePheArgAlaAlaValSerThrArgGly 177
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1. .772
/organism="zea mays"
/organism="zea mays"
/or___type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone="zMnBTa469B124"
/clone="zMnBTa469B124"
/clone="zMnBTa469B124"
/orde="vector: pCR4-TOPO; Site_1: EcoRI; 0.6-1.0 kb high
cor selected genomic DNA library"
a 198 c 216 g 208 t
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  549 ATCTGGAGG--------TCACAGACTTTGCTGACGATGAGATCAAGGGTG 508
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ---GlnAlaPro 103
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 507 ACACAGATGATGACCAGG-----AAGAGCAGGACCCAGTGGTTTCCCTACATGCAATAA 454
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             453 CTGGGATTAGAAGGGAGGACACCATGCAGCTGCGGGTAGTCCTCAAGGGTCAGGAGCTGC 394
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                393 TGGCCCTACTTGACACGGGTTCAACACATAACTTCATCAACTGCAAGGCGGCA-CAACAG 335
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     138 ArgGlySerLeuLeuSerProArgProlleSerTyrLeuLysGlySerSerGlyGlyPro 157
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      663 CAGTCTCGTTCAAGCGCCTCACCCCCGCGGAGTGGCGGAACGACGTAACAAGAGGCTAT 604
                                         Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Bermatophyta; Pagoralotophyta; Esparnatophyta; Pagoralotophyta; Diliopsida; Poales; Poaceae; PacCAD clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 772)
1 (bases 1 to 772)
2 (bases 1 to 772)
3 (bases 1 to 772)
4 (bases 1 to 772)
5 (bases 1 to 772)
6 (bases 1 to 772)
7 (bases 1 to 772)
8 (bases 1 to 7
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    GlnGlyThrGlnLysThrSerHisThrGlyArgAspLysAsnGlnVal ---GluGlyGlu
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       48 ValGlnIleValSerThrAlaThrGlnThrPheLeuAlaThrSerIleAsnGlyVal---
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               -----TyrLeuValThrArgHisAlaAspValIleProValArgArgArgGlyAspSer
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               603 GITACAATIGIGACG-----AGCCCTAIGTACGAGGCCACCGCIGICCCAGGCTAITCI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    67 LeuTrpThrValTyrHisGlyAlaGlyThrArgThrIleAlaSerProLysGlyProVal
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                87 ThrGlnMetTyrThrAsnValAspLysAspLeuValGlyTrp-----
                                                                                                                                                                                                                                                                                                                                                                                                               9712 Medical Center Drive, Rockville, ND 20850,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      772
49
24
67
10
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Matches:
Conservative:
Mismatches:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      US-09-965-594-22 (1-197) x CC406704 (1-772)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                        Tel: 301-838-5843
Fax: 301-838-0208
Email: whitelaw@tigr.org
Seg primer: TR
Class: sheared ends.
                                                                                                                                                                                                                                                                                                                                                 Contact: Cathy Whitelaw
                                                                                                                                                                                                                                                                            Unpublished
Other_GSSs: PURKD12TD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      19
95.50
39.46%
26.49%
9.40%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Percent Similarity:
Best Local Similarity:
                        Zea mays
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               104
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Query Match:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Pred. No.:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   BASE COUNT
                    ORGANISM
                                                                                                                                                                                                                                             TITLE
JOURNAL
COMMENT
                                                                                                                                                    RFERENCE
                                                                                                                                                                                   AUTHORS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             FEATURES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ORIGIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                δλ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ô
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                          /www.rpc-mnun.
/db_xref-*taxon:305*
/db_xref-*taxon:305*
/db_xref-*taxon:305*
/clone_lib-*c. reinhardtii CC-1690, Stress condition 1,
normalized, Lambda Zap II*
/note-*Vector: paluescript II SK-; Site_I: EcoRI: Site_2:
/note-*Vector: paluescript II SK-; Site_I: Cond into lambda land in harman land into lambda land in the EcoRI (5') and XhoRI (3') sites:
/stratagene) in the EcoRI (5') and XhoRI (3') sites:
/stratagene) in the EcoRI (5') and XhoRI (3') sites:
/paluescript II SK- plasmids were cxclsed from the lambda land in library was normalized using method 4 described in Bonaldo et al (1996) Genome Research 6: 791-806.*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CC406704 172 bp DNA linear GSS 19-MAY-2003 PUHKD12TB ZM_0.6_1.0_KB Zea mays genomic clone ZMMBTa469B24, genomic survey sequence. CC406704 CC406704.1 GI:30886794 GSS.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        138
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 LeuLysGlySerSerGlyGlyProLeuLeuCysProAlaGlyHisAlaValGlyllePhe 169
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 171 CACCACCATACCCTTGCTCTCACGTTGCTCACCAAAATTATGCCCATACGGGCCACTA 230
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ThrGlnMetTyrThrAsnValAspLysAspLeuValGlyTrpGlnAlaProGlnGlySer 106
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    284
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ArgSerLcuThrProCysThrCysGlySerSerAspLcuTyrLcuValThrArgHisAla 126
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           345 GAGATCATTTGCATGTGGCTTTAGTCACCCCAAGAAGAGCCTGGGAGTGGGCATTTATAA 404
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   GAAGGGGACGGGAATTCCGTTTGCGGAAAGTGAGCGCCCCCAAGGTCTGACCAAGTGCTA 464
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                525 AAAGAAACCATTTGAGTAGGAAGTCGCCGTTTACCCCCGAAGGTGAAGGTCACCTCTAT 584
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ----GlyProVal 86
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     -----GlySerLeuLeuSerProArgProlleSer---Tyr
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    231 ACAAAGTTACATACACGG------AAGGACCAGCGCGCTTGGCCCCCCCTTGGAGCCG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     AspValIle------ProValArgArgArgGlyAspSerArg-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CTCCAAGGCAGCAATGGGAGCCTTTCGCGGGTGTGCGGGTGCTCGTCCTCTAATGTCAGG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ValAspPheIleProValGluSerLeuGluThrThrMetArg 195
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 GTGGACCGCATACATCTGGAAGACACACGGTGCACTCTACGA 626
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   TyrHisGlyAlaGlyThrArgThrIleAlaSerProLys-----
/organism-"Chlamydomonas reinhardtii"
/mol_type="mRNA"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            701
38
22
63
63
7
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Length:
Matches:
Conservative:
Mismatches:
Indels:
Gaps:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     US-09-965-594-22 (1-197) x BF863244 (1-701)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        16.8
95.50
38.96%
24.68%
9.40%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Best Local Similarity:
Query Match:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         mays
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            173
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Percent Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Zea
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      7.1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   285
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 139
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 150
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  465
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              585
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      81
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     101
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   182
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 127
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CC406704/c
LOCUS
DEFINITION
```

ACCESSION VERSION KEYWORDS SOURCE

8

õ

Qy 104 GlnGlySerArgSerLeuThrProCysThrCysGlySerSerAspLeu119 Db 320 CTGGGATTAGAAGGGAGACACGAGGAGTACTCCGAAAGGGAGGAGGAGAAGGGAAAGGGAAAGGGAAAGGGAAAGGGAAAGGGAAAGGGAAAGGGAAAGGGAAAGGGAAAGGGAAAGGGAAAGAG	120TyriceuvalthrargHisalaAspvallieProValargArgArgGiyAspSer	138 ArgGlySerLeuLeuSerProArgProlleSerTyrLeuLysGlySerSerGlyGlyPro	Oy 158 LeuLeuCysProAlaGlyHisAlaValGly1lePheArgAlaAlaValSerThrArgGly 177 ::	Oy 178 ValAlaLysAlaValAspPhe	Oy 187 ValGluSerLeuGlu 191 ::::::::        Db 580 TTGGATACATTTGAG 594	RESULT 10 BF304699/C LOCUS BF304699 GF 11-NOV-2000 DEFINITION 601888952F1 NTH MGC 17 Home samiens chan clone IMARE-412276 5,		SM HOM Euk Mam	REFERENCE 1 (bases 1 to 984) AUTHORS NIH-MGC http://mgc.nci.nih.gov/. TITLE National Institutes of Health, Mammalian Gene Collection (MGC) JOHNAN. Unmulhished		found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov Plate: LLCMIOO5 row: 9 column: 13 High quality sequence stop: 646. FEATURES Location/Qualifiers 1. 984 //organism="Homo sapiens"	/db_xref="taxon:9606" /clone="Index:412227" /+cone="twom"twom="twom="twom"twom="twom="twom"twom="twom="twom"twom="twom"twom="twom"twom="twom="twom"twom="twom="twom"twom="twom="twom"twom="twom"twom="twom"twom="twom"twom="twom"twom="twom"twom="twom"twom="twom"twom="twom="twom="twom"twom="twom"twom="twom"twom="twom"twom="twom"twom="twom"twom"twom="twom"twom"twom="twom"twom"twom="twom"twom"twom="twom"twom"twom="twom"twom"twom"twom="twom"twom"twom"twom"twom"twom="twom"twom"twom"twom"twom"twom"twom"twom	/tisbue_rype_inabdumiyosarcoma /lab_host="bHOB" (phage_resistant)" /clone_lib="NIH_MGC_l7" /note="organ muscle; Vector; porB7; Site_l: EcoR1;	Since_1: Annul; CDNA Madde by Oilyo'ul Primining the Directionally cloud into EccRI, Ahol sites using the following 5' adaptor: GGCACGAG(G). Size-selected >500bp for average insert size 1.8Mb. Library constructed by the average insert size 1.8Mb. Library constructed by	OUNT 13	Alignment Scores: 26.1 Length: 984
Oy 178 ValAlaLysAlaValAspPhe	Qy 187 ValGluSerLeuGlu 191 ::::::::         Db 193 TTGGATACATTTGAG 179	RESULT 9 CC406705 LOCUS CC406705 CC406705 DEFINITION PUHKD12TD ZM_0.6_1.0_KB Zea mays genomic clone ZMJBTa469B24,	10N CC406705 BY CC406705.1 GI:30886795 DS GSS.	SOURCE  Sea mays  ORGANISM 2ea mays  Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  Spermatophyta; Magnollophyta; Llllopsida; Poales; Poaceae; PACCAD	REFERENCE 1 (bases 1 to 789) AUTHORS Whitelaw.C.A., Quackenbush.J., Van Aken.S., Utterback,T., Resnick A., Fraser.C.M., Yuan,Y., San Miguel.P., Ma.J. and Bennetzen.J. TITLE MAze Genomics Consortium	A.L.	9712 Medical Center Drive, Rockville, MD 20850, USA Tel: 301-838-5843 Fax: 301-838-0208 Email: whitelaw@tigr.org	Seq primer: TF Class: sheared ends. Location/Qualifiers Source 1.789	/organism="2ea mays" /mol_type="genomic DNA" /strain="87" //h vrof="taxon 4577"	/clone-"znwBra69B24" /clone_lib-"zm_0.6_1.0_KB" //nohe="vector: pCR4 TOPO; Site_1: BcoRI; 0.6-1.0 kb high COT selected genomic DNA library" DRASE COUNT 210 a 220 c 203 g 156 t	Alignment Scores: Pred. No.: Pred. No.: Pred. No.: 95.50 Matches: Percent Similarity: 39.46% Mismatches: 67 Ouery Match: 29.40% Mismatches: 67 Ouery Match: 29.40% Mismatches: 46 DB: 29.40% Mismatches:		Oy 29 GlnGlyThrGlnLysThrScrHisThrGlyArgAsplysAsnGlnValGluGlyGlu 47	Qy 48 ValGln1leValSerThrAlaThrGlnThrPheLeuAlaThrSerIleAsnGlyVal 66	Oy 67 LeutrpThrValTyrHisGlyAlaGlyThrArgThrIleAlaSerProLysGlyProVal 86 ::::   ::: ::: :::	Qy 87 ThrGlnMetTyrThrAsnValAspLysAspLeuValGlyTrpGlnAlaPro 103

end repaired, adaptor ligated and size fractionated using sephadex. The resulting fragments were between 0.8 and 3 kb and were cloned into the vector (.x/y reads in M13mp19, .b/g reads in pUC19). The same ligation was transformed in either JM107 or DH5a."  BASE COUNT 108 a 251 c 232 g 110 t ORIGIN  Alignment Scores: 18.8 Length: 701  Score: 85.00 Matches: 61  Percent Similarity: 36.11% Mismatches: 70  Best Local Similarity: 28.24% Mismatches: 70	Ouery Match: 9.35% Indels: 69 DB: 29 Gaps: 11 US-09-965-594-22 (1-197) x Bz342381 (1-701) Oy 23 GlnGlnThrArgGlyGluGlnGlyThrGlnLysThrSerHisThrGlyArgAspLys 41		Db   521 TGGGGCAGCCACGCGCCATGCCGTGCGGGGTGTTTTTTGAACGTACCGCAC 462     Qy	Oy 109 LeuthProCysthr	132 Argargaloglyasp	173 ISETTHRAGGLYValAlalysAlaValAspPheilePro 186 173 ISETTHRAGGLYValAlalysAlaValAspPheilePro 186 174
Score:         95.50         Matches:         39           Percent Similarity:         42.86%         Conservative:         6           Best Local Similarity:         37.14%         Mismatches:         26           Query Match:         9.40%         Indels:         34           DB:         10         Gaps:         7           US-09-965-594-22 (1-197) x BF304699 (1-984)         7         7           QY         100 TrpGlnalaProGlnGlySerArgSerLeuThrProCysThrCysGlySerSerAsp 118         11           QY         111 :::   :::        :::                       599           Db         646 TGGCCCAGTCCACGCATTCCCGGAGAAGAGCACCGTGTACCTGC		159 LeuCysProAlaGlyHis-AlavalGlyIlePheArgAlaAlaVa	ਜੋ ਚੋਂ ਹੁਣ	ACCESSION B2342381 KEYWORDS GSS. SOURCE SORghum blocolor (sorghum) ORGANISM Sorghum bicolor (sorghum) Eukaryota; Viridiplantae: Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Lilopsida; Poales; Poaceae; PACCAD	REFERENCE (Jacks 1 to 701) AUTHORS Rabinowicz, P.O. O'Shaughnessy, A.L., Ballja, V., Dedhia, N., Katzenburger, P.D., O'Shaughnessy, A.L., Ballja, V., Dedhia, N., Katzenburger, P.D., O'Shaughnessy, A.L., Ballja, V., Bascimento, L., Zutavern, T., Palmer, L., McCombie, W.R. and Martienssen, R.A.  TITLE Genomic shotgun sequences from Sorghum bicolor (methyl-filtered) JODRNAL Unpublished CONTACT: W. Richard McCombie CONTACT: W. Richard McCombie CONTACT: W. Richard McCombie Cold Spring Harbor Laboratory PO BOX 100, Cold Spring Harbor, NY 11724, USA Fee: 516 367 884 Spring Harbor, NY 11724, USA	Email: mccomble@cshl.org  Plate: 1681 row: b column: 11 Seq primer: -21MillonivFwd Class: shorgun High quality sequence stop: 701.  Location/Qualifiers 1. 701.  /organism="Sorghum bicolor" //mol_Lype="genomic DNA" //db_xref="taxon:458" //db_xref="taxon:458" //dbc="losabli" //dbc="site_1: Xba I; Site_2: Xba I; The vector was digested with XbaI and one nucleotide was added by fill in the recessive 3' end. The genomic DNA was nebulized,

525 TCCCAGGGATACTCTCAGACATGGGCCACCTACACGGATGAA 72 HisGlyAlaGlyThrArgThrIle	DD 55/ CATANCECCGAGGGTCAGAATIGCCACAGAGGGGGGGGGGGGGGGGG	Qy 90TyrThrAsnValAspLysAspLeuValGlyTrpGlnAlaProGlnGly 105	Qy 106 SerArgSerLeuThrProCysThrCysGlySerSerAspLeu 119	Qy 120 TyrLeuValThrArgHisAlaAspValIleProValArgArgArgGlyAspSerArgGly 139	Oy 140 SerLeuLeuSerProArgProlleSerTyrLeuLysGlySerSer 154	Oy 155GlyGlyProLeuLeuCy8ProAlaGlyHis164	Oy 165	RESULT 13 BM467279 LOCUS DEFINITION	ACCESSION BM467279 VERSION BM467279.1 GI:18516321 KEYWORDS EST. SOURCE Home sapiens (human) ORGANISM Home sapiens			High quality sequence Stop: 1/0.  FEATURES  1. 1440  Anotation/Qualifiers  //noganism="Homo sapiens" //mofarism="Homo sapiens" //mofarism="Homo sapiens" //db_xref="taxon:9606" //lone="IMAGE:5505315" //tab_host="DH10B (phage resistant)" //clone_lib="NIH_MGC_85"
Saccharomycetales; Dipodascaceae: Yarrowia. 1 (bases 1 to 1062) Souciet.J.L., Aigle,M., Artiguenave,F., Blandin,G., Bolotin-Fukuhara,M., Bon,E., Brottier,P., Casaregola,S.,	<pre>de-Wontigny,J., Dujon,B., Durrens,P., Lepingle,A., Liorente,B., Malpertuy,A., Neuveglise,C., Ozier-Kalogeropoulos,O., Potier,S., Saurin,W., Tekaia,F., Toffano-Nioche,C., Wesolowski-Louvel,M., Winoker,P. and Weissenbach,J. Genomic exploration of the hemiascomycetous yeasts: 1. A set of</pre>		2 (bases 1 to 1062) Casaregola,S., Neuveglise,C., Lepingle,A., Bon,E., Feynerol,C., Artiguenave,F., Wincker,P. and Gaillardin,C. Genomic exploration of the hemiascomycetous yeasts: 17. Yarrowia	lipolytica FEBS Lett. 487 (1), 95-100 (2000) 20584727 11152892	3 (bases 1 to 1062) Genoscope. Direct Submission Submitted (07.SEP-2000) Genoscope - Centre National de Sequencage,	ref@genoscope.cns.fr	usus, Saccinaromyces servazzii, Zygosaccharomyces rouxii.  charomyces kluyveri, Kluyveromyces thermotolerans, Kluyveromyces  Lis var. lactis, Kluyveromyces marxianus var. marxianus, Pichia  stat, Debastyomyces hansenii var. hansenii, Pichia sorbitophila,  did tropialis and Varrowia libolutica. Genomic inserts of 3 to	5 kb were prepared and both extremities were sequenced. See keywords for description of this sequence and for the sequence of the other extremity of this insert.  Location/Qualifiers  1. 1062	/organism="Yarrowia lipolytica" /mol_type="genomic DNA" /strain="CLIB 89" /db_xref="taxon:4952" /clone="AWOAA006803"	/clone_lib="AWOAA" /note="end : T3*  <594>917 /note="similar to Saccharomyces cerevisiae ORF YKR093w [ PTR2 ; peptide transporter ]* /evidence=not_experimental	36.3 Length: 10 94.50 Matches: 53 25.62% Conservative: 25 ty: 24.20% Mismatches: 62 129 30% Indels: 79 29 Gaps: 12	US-09-965-594-22 (1-197) x CNS06OHN (1-1062)           QY         18 AspThralaTyralaGlnGlnThrArgGlyGluGlnGlyThrGlnLySThrSerHisThr 37           II         ::

```
/dev_stage="adult"
/clone_lib="An expressed sequence tag (EST) collection
/clone_lib="An expressed sequence tag (EST) collection
/clone_lib="An expression plant Selaghinella lepidophylla"
/note="Vector: Lambda Uni.Zap XR. Bluescript SK:: Site_l:
ECORI; Site_2: Xhol; Library construction was performed
according to manufacture's (Stratagene, Inc.) recommended
protocol for the Lambda UniZapXR vector and cDNA synthesis
kit."
                                                                                                                                                                                                                                                                                                                                                                                                                      /tissue_type="microphyll fronds undergoing desiccation for 2.5 h"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           94 AspLysAspLeuValGlyTrpGlnAlaProGlnGlySerArgSerLeuThrProCysThr 113
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              173 GGATTGGATCATACGCTGACCACAGGCGTCATCAGTGGTCTTCGAAGGGAGATTACT--- 229
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1 (Dases 1 to 528)
Iturriaga,G. and Cushman,J.C.
An expressed sequence tag (EST) collection from the resurrection
plant Selaginella lepidophylla
Unpublished
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   GlySerLeuLeuSerProArgProIleSerTyrLeu--------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      151 LysGlySerSerGlyGlyProLeuLeuCysProAlaGlyHisAlaValGlyIlePheArg
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         171 AlaAlaValSerThrArgGlyValAlaLySAlaValAspPhe---IleProValGluSer
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          GACAAGGATGTAGCGGTGCTGAAGATCGATGCTCAAGCAACAGATCTCAGCCCAATACCC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           113 CITGGAAGTICGTCCGATCTGCTTGTTGGCCAGAAGGTGTATGCTATCGGTAATCCTTTT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ------ThrArgHisAlaAspValIleProValArgArgArgGlyAspSerArg
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        287 CCTGGGAACAGCGGGGGTCCGCTATTGGACAGTTCTGGAAATTTGATAGGCATCAACACT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            114 CysGlySerSerAspLeuTyrLeuVal------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  528
37
15
43
43
                                                                                                                                                                                                                                                                                                                                                         /organism="Selaginella lepidophylla"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Length:
Matches:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               137
                                                                                                                                                                                                                                                                                                                                                                     /mol_type="mRNA"
/db_xref="taxon:59777"
/clone="SLA005F12"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          US-09-965-594-22 (1-197) x BM402566 (1-528)
                                                                                                                                                 USA
                                                                                                                                                                                                                                                                     column: 12
                                                                                          Contact: Cushman JC
Department of Biochemistry
University of Nevada
MS200, Reno, NV 89557-0014, U.
Tel: 715-784-1918
Fax: 775-784-1650
Email: jcushman@unr.edu
                                                                                                                                                                                                                                                                Plate: 005 row: F column: 1
Seg primer: T3 20mer
High quality sequence stop: 5
Location/Qualifiers
1. 528
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 137
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                20.5
93.00
42.98%
30.58%
9.15%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               125 c
                                                                                                                                                                                                                                  FORWARD: T3 20mer
BACKWARD: T7 21mer
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Percent Similarity:
Best Local Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Leu 190
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 129
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Alignment Scores:
Pred. No.:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   230
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              190
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 139
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              123
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Query Match:
DB:
                                                                                                                                                                                                                                                                                                                                          source
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               BASE COUNT
ORIGIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        :
2
                           AUTHORS
TITLE
                                                                              JOURNAL
                                                                                                                                                                                                                                                                                                                        FEATURES
                                                                                              COMMENT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            οχ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 BM402566 528 bp mRNA linear EST 01-JUL-2002 SAB005F12.34513 An expressed sequence tag (EST) collection from the resurrection plant Selagainella lepidophylla Selaginella lepidophylla cDNA clone SLA005F12 5, mRNA sequence.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1083 CCACACCCGCATTTTACTCCCCGCTATATTCTNTTATCTGNTGCGGNCGGACGCCCCGC 1142
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        149
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Selaginella lepidophylla
Selaginella lepidophylla
Sukaryota, viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Lycopodiophyta; Isoetopsida; Selaginellales; Selaginellaceae;
Selaginella.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 116
                                                                                                                                                                                                                                                                                                                                                                                                                                          920
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             980
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 117 -------SerAspLeuTyrLeuValThrArgHisAlaAspValIleProVa 131
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         149 rLeuLysGlySerSerGlyGlyProLeuLeuCysProAlaGlyHisAlaValGlyIlePh 169
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            eArg-AlaAlaValSerThrArgGlyValAlaLySAlaValAspPheIleProValGluS 189
                                                                                                                                                                                                                                                                                                      753 CAAATCACCATATCCGGAGATGTTGCTCTCTGT------TTTAAGCGTCACACA 800
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            85 oValThrGlnMetTyrThrAsnValAspLysAspLeuValGlyTrp------ 100
                                                                                                                                                                                                                                                                                                                                                                                                          89
                                                                                                                                                                                                                                                                                                                                                                                                                                                                           85
                                                                                                                                                                                                                                                                                                                                      49
         full-length clones and constructed by Life Technologies. Note: this is a NIH_MGC Library."

442 c 308 g 355 t 5 others
                                                                                                                                                                                                                                                                                                                                                              301 CCCGCGACCAACTCTACTCCCCGCCACAAAATATACTTTCGTGGAGCGGGAATATCTT
                                                                                                                                                                                                                                                                                                                                                                                                                            -----GlnAlaProGlnGlySerArgSerLeuThrProCysThrCysGlySer--
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          12 ArgileAsnLeuSerGlyAspThrAlaTyrAlaGlnGLnThrArgGlyGluGlnGlyThr
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            921 ATTITITITICCCTGGCAGGCACCTCTCGAATTCCCGAGCGCGCGGTTCCCTCGCTCC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      131 lArgArgArgGlyAspSer-----ArgGlySerLeuLeuSerProArgProlleSerTy
                                                                                                                                                                                                                                                                                                                                      32 GlnLysThrSerHisThrGlyArgAspLysAsnGln-----ValGlu-GlyGluValGl
                                                                                                                                                                                                                                                                                                                                                                                                          49 nIleValSerThrAlaThrGlnThrPheLeu---AlaThrSerIleAsnGlyValLeuTr
                                                                                                                                                                                                                                                                                                                                                                                                                                                                           68 pThrValTyrHisGlyAlaGlyThrArgThrIleAlaSerProLysGly-----Pr
                                                                                                            1440
54
22
95
35
                                                                                                            Length:
Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                    US-09-965-594-22 (1-197) x BM467279 (1-1440)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       BM402566.1 GI:21643782
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1320 TICTCGTCACCACA 1333
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              erLeuGluThrThr 193
                                                                                                            54.2
94.50
36.89%
26.21%
9.30%
                                                                                                                                                                   Similarity:
                                               ø
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      BM402566
                                                                                                                                                 Percent Similarity:
                                                                                            Alignment Scores:
Pred. No.:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            169
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 101
                                                                                                                                                                              Query Match:
DB:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     VERSION
KEYWORDS
SOURCE
ORGANISM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    DEFINITION
                                                                                                                                                                   Best Local
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 RESULT 14
BM402566
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ACCESSION
                                             BASE COUNT
                                                               NIBING
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              윱
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                  a
                                                                                                                                                                                                                                                                                                                                                                                                        à
                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          a
a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
```

150 286 346

170

---ArgSerLe 109

```
129 eProValArgArgArgGlyAspSerArgGlySerLeuLeuSerProArgProlleSerTy 149
                                                                                                                                                                                                                                        149 rLeuLysGlySer-SerGlyGlyProLeuLeuCysProAlaGlyHisAlaValGlyIleP 169
                                                                                109 uThrProCysThrCysGlySerSerAspLeuTyrLeuValThrArgHisAlaAspValI1 129
                                                                                                                       -----CTGGAGCAGAT 277
                                                                                                                                                                                                   278 ATGACGAGCCATGGAGGGAGAGAGCAGGAGCCTTCTCCCTGGCCCAGCCCCTGGACCATÀ 337
                                                                                                                                                                                                                                                                               197 AGAGCTTGTGTGCTGCCAACACCCCCAGGGAGAACTTGCTGCATGGGGGAGGAGGTCTGT
    95 sAspLeuValGlyTrpGlnAlaProGlnGlySer---
                                                                                                                                                                                                                                                                                                                                                                                                                    Search completed: August 31, 2003, 04:27:50 Job time : 1914.31 secs
                                                                                                                     257 ACATAGATGT
                                                                                                                                                                                                                                                                                                                      169 heArgAlaAla 172
                                                                                                                                                                                                                                                                                                                                                             389 TCACAGGGGT 399
                                                                                                          QQ
                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                      a
                                                                                                                                                                                                                                                                                                                    ò
                                         d
                                                                                ò
                                                                                                                                                     ò
                                                                                                                                                                                                                                     ò
                                                                           AQ538021 550 bp DNA linear GSS 18-MAY-1999
RPCI-11-32014.TJ RPCI-11 Homo sapiens genomic clone RPCI-11-32014,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     113 CAAGAG------CACGGATTTGCAAAATGTGG------CACGGATTTGGA 145
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     53 CACCCTGGAAGGCTATGAGATCATAGGAGGCTTCTGCCAGTCTCTGTCATCCACTGGC 112
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          library availability, please contact Pieter de Jong (pleter@dejong.med.buffalo.edu). Clones may be purchased from BACPAC Resources (http://Dacpac.med.buffalo.edu/ordering) or from Research Genet cs (info@resgen.com). BAC end search page: http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html. Seq primer: SP6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              56 GlnThrPheLeuAlaThrSerlleAsnGlyValLeuTrpThrValTyrHisGlyAlaGly 75
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        76 ThrArgThrIleAlaSerProLysGlyProValThrGlnMetTyrThrAsn-ValAspLy 95
                                                                                                                                                                                                                 Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Hammalla; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (Dases 1 to 560)
Zhao,S., Adams,M.D., Nierman,W., Malek,J., de Jong,P. and Venter
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                HisThrGlyArgAspLysAsnGlnValGluGlyGluValGlnIleValSerThrAlaThr 55
                                                                                                                                                                                                                                                                                                                                    Use of BAC End Sequences from Library RPCI-11 for Sequence Ready
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      /db_xref="taxon:9606"
/db_xref="taxon:9606"
/cel="taxon:9606"
/sex="Male"
/cel="type="Lymphocytes"
/clone_lib="ReCi=1"
/note="Vector: pBACe3.6; Site_1: EcoRI; Site_2: EcoRI;
RPCI11 Human Male BAC Library"
a 125 c 181 g 124 t
                                                                                                                                                                                                                                                                                                                                                                                           Other_GSS: RPCI-11-32014.TV
Contact: Shaying Zhao, William Nierman, Mark Adams
Contact: Shaying Zhao, William Nierman, Mark Adams
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850
Tel: 301 838 0200
Fax: 301 838 0200
Email: hbe@tigr.org
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   560
44
11
54
36
6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Length:
Matches:
Conservative:
Mismatches:
Indels:
Gaps:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    /organism-"Homo sapiens"
/mol_type-"genomic DNA"
/db_xref-"GDB:7622691"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           US-09-965-594-22 (1-197) x AQ538021 (1-560)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Location/Qualifiers
1. .560
                                                                                                                       genomic survey sequence
                                                                                                                                                     AQ538021.1 GI:4849711
GSS.
                                                                                                                                                                                                 Homo sapiens (human)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 22.1
93.00
37.93%
30.34%
9.15%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 BAC ends.
                                                                                                                                                                                                                                                                                                                                                             Map Building
                                                                                                                                                                                                                                                                                                                                                                               Unpublished
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Percent Similarity:
Best Local Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ಸ
GTT 409
                                                                                                                                       AQ53802]
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       130
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Alignment Scores:
Pred. No.:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                36
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Query Match:
DB:
                                                                                                                                                          VERSION
KEYWORDS
SOURCE
ORGANISM
                                                                                                DEFINITION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     BASE COUNT
ORIGIN
                                                                                                                                                                                                                                                                          REFERENCE
AUTHORS
                                                                                                                                       ACCESSION
                                                                                                                                                                                                                                                                                                                                                                             JOURNAL
                                         RESULT 15
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               FEATURES
                                                             AQ538021
                                                                                                                                                                                                                                                                                                                                      TITLE
                                                                                                                                                                                                                                                                                                                                                                                               COMMENT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             qq
```

```
Hepatitis C virus
Hepatitis C virus
Hepatitis C virus
Synthetic nonstruc
Peptide KKNS4a use
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          KKNS4a peptide. S
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Hepatitis C virus
Peptide KKNS4a for
                                                                                                                                    August 30, 2003, 17:42:58 ; Search time 2.49162 Seconds (without alignments) 700.745 Million cell updates/sec
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1: /SIDSI/gcgdata/geneseq/geneseqp-embl/AA1980.DAT:*
2: /SIDSI/gcgdata/geneseq/geneseqp-embl/AA1981.DAT:*
3: /SIDSI/gcgdata/geneseq-embl/AA1981.DAT:*
4: /SIDSI/gcgdata/geneseq-embl/AA1982.DAT:*
5: ISDISI/gcgdata/geneseq-embl/AA1983.DAT:*
5: SIDSI/gcgdata/geneseq-embl/AA1986.DAT:*
6: /SIDSI/gcgdata/geneseq-embl/AA1986.DAT:*
7: /SIDSI/gcgdata/geneseq-embl/AA1986.DAT:*
7: /SIDSI/gcgdata/geneseq-embl/AA1986.DAT:*
8: /SIDSI/gcgdata/geneseq-embl/AA1980.DAT:*
9: /SIDSI/gcgdata/geneseq-embl/AA1980.DAT:*
11: /SIDSI/gcgdata/geneseq-embl/AA1991.DAT:*
12: /SIDSI/gcgdata/geneseq-embl/AA1991.DAT:*
13: /SIDSI/gcgdata/geneseq-embl/AA1991.DAT:*
14: /SIDSI/gcgdata/geneseq-embl/AA1991.DAT:*
15: /SIDSI/gcgdata/geneseq-embl/AA1991.DAT:*
16: /SIDSI/gcgdata/geneseq-embl/AA1991.DAT:*
17: /SIDSI/gcgdata/geneseq-embl/AA1991.DAT:*
18: /SIDSI/gcgdata/geneseq-embl/AA1991.DAT:*
18: /SIDSI/gcgdata/geneseq-geneseq-embl/AA1995.DAT:*
18: /SIDSI/gcgdata/geneseq-geneseq-embl/AA1995.DAT:*
18: /SIDSI/gcgdata/geneseq-geneseq-embl/AA1995.DAT:*
19: /SIDSI/gcgdata/geneseq-geneseq-embl/AA1995.DAT:*
20: /SIDSI/gcgdata/geneseq-geneseq-embl/AA1997.DAT:*
21: /SIDSI/gcgdata/geneseq-geneseq-embl/AA1997.DAT:*
22: /SIDSI/gcgdata/geneseq-geneseq-embl/AA1997.DAT:*
23: /SIDSI/gcgdata/geneseq-geneseq-embl/AA1997.DAT:*
24: /SIDSI/gcgdata/geneseq-geneseq-embl/AA1997.DAT:*
25: /SIDSI/gcgdata/geneseq-geneseq-embl/AA1997.DAT:*
26: /SIDSI/gcgdata/geneseq-geneseq-embl/AA1997.DAT:*
27: /SIDSI/gcgdata/geneseq-geneseq-embl/AA1997.DAT:*
28: /SIDSI/gcgdata/geneseq-geneseq-embl/AA1997.DAT:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA2000.DAT:*/SIDS1/gcgdata/geneseq/geneseqp-emb1/AA2001.DAT:*/SIDS1/gcgdata/geneseq/geneseqp-emb1/AA2001.DAT:*/SIDS1/gcgdata/geneseq/geneseqp-emb1/AA2002.DAT:*/SIDS1/gcgdata/geneseq/geneseqp-emb1/AA2003.DAT:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Description
                                                                                                                                                                                                                                                                                                                                                                                                                                              1107863
GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.
                                                                                                                                                                                                                                                                                                                                                                                                                                            Total number of hits satisfying chosen parameters:
                                                                                                                                                                                                                                                                                                                                                                                               1107863 segs, 158726573 residues
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SUMMARIES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   AAB96864
AAW40552
AAM52536
AAG64158
AAB67391
AAB66371
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         AAB15227
AAM48242
                                                                                             using sw model
                                                                                                                                                                                                                                                                                                                          BLOSUM62
Gapop 10.0 , Gapext 0.5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     A_Geneseq_19Jun03:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Π
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Minimum DB seq length: 0
Maximum DB seq length: 2000000000
                                                                                                                                                                                                                        us-09-965-594-26
49
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1 GSVVIVGRIVL 11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Query
Match Length DB
                                                                                             OM protein - protein search,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    100.0
100.0
100.0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           100.0
100.0
100.0
100.0
100.0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Score
                                                                                                                                                                                                                                  Title:
Perfect score:
                                                                                                                                                                                                                                                                                                                          Scoring table:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Sequence:
                                                                                                                                                                                                                                                                                                                                                                                               Searched:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Database
                                                                                                                                            Run on:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Result
```

### ALIGNMENTS

Modified hepatitis C virus (HcV) NS3 protease comprising at least 1 substitution of a hydrophobic alpha-helix 0 amino acid to a hydrophilic amino acid, useful for screening inhibitors that may treat hepatitis C Hepatitis; NS3 protease; viral replication; chronic liver disease; liver failure; liver cancer. Wittekind M, Weinheimer S, Zhang Y, Goldfarb V; Hepatitis C virus NS4A residues 21-31. AAB15227 standard; protein; 11 AA. (BRIM ) BRISTOL-MYERS SQUIBB CO. 06-JAN-2000; 2000WO-US00345. 99US-0115271. 19-DEC-2000 (first entry) WPI; 2000-465976/40. Hepatitis C virus WO200040707-A1. 08-JAN-1999; 13-JUL-2000. AAB15227 RESULT 1 PART SERVICE S

invention

```
;
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        chimeric Hepatitis C virus (HCV)-pestivirus genome. The construct comprises a pestivirus genome where a Npro pestivirus protease gene is replaced with a gene encoding a functional HCV NS3 protease. Furthermore, each junction site recognised by the Npro protease is replaced with a junction site recognised by the HCV NS3 protease. The construct is useful for screening compounds that inhibit HCV in vivo by inhibiting HCV protease, where screening may be in cell culture or in an animal model. The present sequence is a peptide from the tethered NS4A cofactor sequence from HCV NS3, which was used to illustrate the present
                                                                 The present sequence comprises residues 21-31 of the Hepatitis C virus (HCV) NS4A protease enzyme. It was used to create a number of fusion proteins also containing the HCV NS3 protease. The NS4A and NS3 proteins are both essential for the replication of the virus, acting to cleave its replicative proteins from the polyprotein produced from the HCV genome. Inhibitors of the two proteins should be effective as antiviral treatments of HCV infection. This is useful as HCV can lead to chronic liver disease such as cirrhosis, liver failure and liver cancer. The present invention concerns a number of NS3 mutants and NS3-NS4A fusion proteins which can be used to identify inhibitors of this type, as well
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Nucleic acid construct encoding chimeric Hepatitis C Virus (HCV) pestivirus genome where the Npro protease gene is replaced with NS3 protease gene, useful for in vivo screening of compounds which inhibit HCV infection
                                                                                                                                                                                                                                                                                                                                                                                      Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            The present invention relates to a nucleic acid construct encoding a
                                                                                                                                                                                                                                                                                                                                                                                    ;
                                                                                                                                                                                                                                                                                                                                              100.0%; Score 49; DB 21; Length 11; 100.0%; Pred. No. 0.025; 1ve 0; Mismatches 0; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Pestivirus; Npro; protease; NS3; screening; NS4A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Hepatitis C virus NS4A cofactor peptide.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  AAM48242 standard: Peptide: 12 AA.
                                    Claim 15; Page 22; 66pp; English
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Example 1; Fig 2; 20pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               99US-0344456
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               99US-0344456.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Hong Z, Lai VCH, Lau JYN;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (first entry)
                                                                                                                                                                                                                                                                                                                                                Query Match 100.

Best Local Similarity 100.
Matches 11; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                       1 GSVVIVGRIVL 11
                                                                                                                                                                                                                                                                                                                                                                                                                                          GSVVIVGRIVE 11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (SCHE ) SCHERING CORP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           WPI; 2002-121103/16.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Hepatitis C virus
                                                                                                                                                                                                                                                                                                               11 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             25-JUN-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               25-JUN-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      JS6326137-B1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         21-MAR-2002
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        34 - DEC - 2001
                                                                                                                                                                                                                                                                               complexes
                                                                                                                                                                                                                                                                                                               Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     AAM48242;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               RESULT 2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                         ò
```

```
ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ó
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              The present invention describes methods and compositions capable of preventing the replication of hepatitis C virus (HCV), involving administering a compound which inhibits NS2/3 autocleavage. Also provided are peptides capable of inhibiting this cleavage step, of which this sequence is an example. These are useful in the treatment of HCV infection in humans and chimpenzees, and in research applications, for example in studying the stabilisation of NS2/3, the effects of NS2/3 on HCV polyprotein processing and the effects of inhibiting NS2/3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 cell, or
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Gaps
                                                         Gaps
                                                                                                                                                                                                                                                                                             Hepatitis C virus; HCV; NS2; NS3; inhibitory peptide; cleavage; replication inhibition; chimpanzee; human; infection; gene therapy
                                                                                                                                                                                                                                                                 Hepatitis C virus NS2/3 cleavage inhibitory peptide SEQ ID NO: 22
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Inhibiting hepatitis C virus (HCV) replication in HCV infected \alpha in a patient or treating a patient for HCV infection comprises inhibiting autocleavage of NS2/3 .
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ö
                                                         .;
0
                            Length 12;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     100.0%; Score 49; DB 22; Length 13; 100.0%; Pred. No. 0.03;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ö
                       100.0%; Score 49; DB 23;
100.0%; Pred. No. 0.027;
ive 0; Mismatches 0;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Disclosure; Page 46; 50pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               AAW40552 standard; peptide; 23 AA.
                                                                                                                                                                                 AAB96864 standard; peptide; 13 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 0;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ÿ
                                                                                                                                                                                                                                                                                                                                                                                                                        25-AUG-2000; 2000WO-US23444.
                                                                                                                                                                                                                                                                                                                                                                                                                                                  99US-0151395
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Kuo
                                                                                                                                                                                                                                      06-JUL-2001 (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 11; Conservative
                                                        Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1 GSVVIVGRIVL 11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1 GSWVIVGRIVL 11
                                                                                                 1 GSWIVGRIVE 11
                                                                                 1 GSVVIVGRIVL 11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (MERI ) MERCK & CO INC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Darke PL, Jacobs AR,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    WPI; 2001-343059/36.
                         Query Match
Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Sest Local Similarity
                                                                                                                                                                                                                                                                                                                                      Hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         13 AA;
12 AA;
                                                                                                                                                                                                                                                                                                                                                                 WO200116379-A1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                  30-AUG-1999;
                                                                                                                                                                                                                                                                                                                                                                                            08-MAR-2001
                                                        11;
                                                                                                                                                                                                            AAB96864;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Sequence
Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Matches
                                                        Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 RESULT 4
AAW40552
ID AAW4
XX
                                                                                                                                                     RESULT 3
                                                                                                                                                                  AAB96864
                                                                                                                                                                                              ò
                                                                                                           Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ŝ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        a
```

AAW40552;

19-MAR-1998

FOX T,

NAMES OF THE PROPERTY OF THE P

```
The present invention relates to a novel pyrrolopyrazinone derivative, its stereoisomer or salt. It was found that the derivative is a Hepatitis C virus (HCV) nonstructural (NS) 3 protease inhibitor. The derivative is useful for the manufacture of a medicament for the treatment of HCV and in therapy for treating HCV infection. The present peptide was used to illustrate the present invention.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              The invention relates to novel ketoamide and ketoester derivatives for use as inhibitors of hepatitis C virus (HCV) NS3 protease inhibitors. The compounds are useful for treating viral infections e.g. hepatitis C virus. The present sequence was used in an experiment measuring the effect of an inhibitor on the rate of hydrolysis of an ester substrate.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   New ketoamide derivatives useful for treating infections e.g. hepatitis
                                                                                                                                                                                       New pyrrolopyrazinone derivatives useful for treating Hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Hepatitis C virus, HCV, NS3 protease; alpha-ketoamide inhibitor;
virucide; hepatotropic; antiinflammatory; viral infection; KKNS4a.
                                                                                                                                                                                                                                                                                                                                                                                                                                                          Length 23;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                        100.0%; Score 49; DB 22; 1
100.0%; Pred. No. 0.055;
.ive 0; Mismatches 0;
                                                                                                                                                                                                         infection are NS3 protease inhibitors
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Disclosure; Page 195; 282pp; English.
                                                                                                                                                                                                                                             Disclosure; Page 130; 191pp; English
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      AAG64158 standard; peptide; 23 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           01-DEC-2000; 2000WO-US32677
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 990S-0168998
                                     29-FEB-2000; 2000US-185618P.
28-FER-2001; 2001WO-US06269
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (DUPO ) DU PONT PHARM CO.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Best Local Similarity 100.
Matches 11; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1 GSVVIVGRIVL 11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          GSVVIVGRIVE
                                                                         DUPO ) DUPONT PHARM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           WPI; 2001-464936/50.
                                                                                                                                                   WPI; 2001-656752/75
                                                                                                                                                                                                                                                                                                                                                                                                                       23 AA;
                                                                                                             Zhang X, Han W;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 W0200140262-A1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    KKNS4a peptide
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 03-DEC-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            19-OCT-2001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     07-JUN-2001.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Synthetic
                                                                                                                                                                                                                                                                                                                                                                                                                         Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          AAG64158;
                                                                                                                                                                                                                                                                                                                                                                                                                                                          Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Han W;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  RESULT 6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    AAG64158
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            q
ő
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          This is the amino acid sequence of the novel sNS4A (synthetic Hepatitis C nonstructural protein 4A) peptide. It acts as a cofactor for the NS3 protease in order to achieve proteclytic processing of Hepatitis C virus (HCV) nonstructural proteins. It is used in the method of the invention as part of a device which can be used to provide information for the design of drugs for the treatment of HCV infection. They can also be used for determining the 3-dimensional structure of molecules or molecular complexes which contain at least some structurally similar features to a HCV NS3 serine protease domain.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Virucidė: pyrrolopyrazinone derivative; Hepatitis C virus inhibitor; nonstructural 3 protease; NS3 protease; viral infection.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                New hepatitis C virus crystal compositions - comprising a HCV NS3-like polypeptide complexed with a NS4A-like polypeptide, used particularly for drug design
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ..
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  100.0%; Score 49; DB 19; Length 23; 100.0%; Pred. No. 0.055;
                                                                                                         Synthetic Hepatitis C nonstructural protein; sNS4A peptide; cofactor; NS3 protease.
                                                                                                                                                                                                                                                                                                                                                                                                                                        Thomson JA;
                                                                                                                                                                                                                                                                                                                                                                                                                                        Morgenstern KA,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Peptide KKNS4a used in an enzyme assay.
                                                                           Synthetic nonstructural peptide sNS4A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     AAM52536 standard; peptide; 23 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Claim 4; Page 30; 97pp; English.
                                                                                                                                                                                                                                                                                                         97WO-US16182
                                                                                                                                                                                                                                                                                                                                           960S-0731336
960S-0025274
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               (first entry)
                                       (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Ouery Match
Best Local Similarity 100.
Marches 11; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                     (VERT-) VERTEX PHARM INC.
                                                                                                                                                                                                                                                                                                                                                                                                                                      Lin C,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1 GSVVIVGRIVL 11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         WPI; 1998-250953/22.
                                                                                                                                                                       Synthetic.
Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      23 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                          Kim JL,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      WO200164678-A2.
                                                                                                                                                                                                                                                                                                       12-SEP-1997;
                                     20-NOV-1998
                                                                                                                                                                                                                             W09811134-A1
                                                                                                                                                                                                                                                                                                                                                               12-SEP-1996;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               31-JAN-2002
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              07-SEP-2001.
```

Sequence

Synthetic

AAM52536;

RESULT 5

9 ô

ö

Gaps

;

Forsyth TP;

```
The present invention provides a number of boronic acid derivatives which act as inhibitors of the hepatitis C virus NS3 protease enzyme. They can be used to treat infection by the virus, which can cause liver cirrhosis and liver cancer.
                                                                                                                                                                                  New boronic acid derivatives, optionally containing peptides, used to treat hepatitis C infections, are hepatitis C viral protease inhibitors
                                                                                                                                                                                                                        Example 60; Page 208; 258pp; English.
liver cancer; NS3; antiviral agent.
                                                                                                                                          Kettner CA, Jagannathan S,
                                                                                07-JUL-2000; 2000WO-US18655.
                                                                                                                       (DUPO ) DU PONT PHARM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Douty BD,
                                                                                                                                                             WPI; 2001-103001/11.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              WPI; 2002-627251/67
                                                                                                                                                                                                                                                                                              23 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        WO200248116-A2.
                                        W0200102424-A2
                     Unidentified.
                                                                                                   07-JUL-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           10-JUN-2002.
                                                            11-JAN-2001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Glunz PW,
                                                                                                                                                                                                                                                                                                                                                                                                                                                 ABG3250B;
                                                                                                                                                                                                                                                                                               Seguence
                                                                                                                                                                                                                                                                                                                                                                                                         RESULT 9
                                                                                                                                                                                                                                                                                                                                                                                                                                        셤
 ö
                                        ô
                                                                                                                                                                                                                                                                                                                                                                                                                                               The present invention relates to Lactam derivatives. These derivatives may be used for treating hepatitis C virus (HCV) infection. They can also be used for inhibiting HCV in a body fluid sample and as a standard or reagent in a test or assay for determining the ability of a potential pharmaceutical to inhibit HCV NS3 protease and/or HCV growth,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Hepatitis C virus; protease; boronic acid; inhibitor; liver cirrhosis;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Gaps
                                                                                                                                                                                                                                                                                                                                                                                              New lactam derivatives are hepatitis C virus NS3 protease inhibitors useful for treating HCV infections
                                         Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ö
                                        ;
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        100.0%; Score 49; DB 22; Length 23; 100.0%; Pred. No. 0.055; ive 0; Mismatches 0; Indels
                   100.0%; Score 49; DB 22; Length 23; Ilarity 100.0%; Pred. No. 0.055; Conservative 0; Mismatches 0; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Hepatitis C virus protease inhibitor related peptide #1.
                                                                                                                                                                                                                 Lactam; hepatitis C virus; HCV; NS3 protease.
                                                                                                                                                                                                                                                                                                                                                                                                                             Example 26; Page 100; 130pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       AAB66371 standard; peptide; 23 AA.
                                                                                                                                 AAB67391 standard; peptide; 23 AA.
                                                                                                                                                                                                                                                                                               26-JUL-2000; 2000WO-US20189
                                                                                                                                                                                                                                                                                                                 99US-0145631
                                                                                                                                                                                                                                                                                                                                                          Decicco CP;
                                                                                                                                                                         26-APR-2001 (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Local Similarity 100.
                                                                                                                                                                                                                                                                                                                                     (DUPO ) DU PONT PHARM CO.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1 GSVVIVGRIVL 11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            WPI; 2001-159696/16.
                                Local Similarity
nes 11; Conser
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     23 AA;
  Š
                                                                                                                                                                                            KKNS4a peptide
                                                                                                                                                                                                                                                      WO200107407-A1
                                                                                                                                                                                                                                                                                                                                                         Priestley ES,
 23
                                                                                                                                                                                                                                                                                                                 26-JUL-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              09-APR-2001
                                                                                                                                                                                                                                                                           01-FEB-2001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Query Match
Best Local S
                                                                                                                                                                                                                                    Synthetic
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          AAB66371;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Sequence
                                                                                                                                                     AAB67391;
  Sequence
                       Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Matches
                                         Matches
                                                                                                              RESULT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          XX DXX
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Dp
  S
                                                                                g
```

```
New pyrimidinones useful as serine protease inhibitors in the treatment of e.g. viral infection
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               NS3: HCV; protease; HCV infection; hepatitis; cirrhosis; liver cancer; pyrimidinone; serine protease inhibitor; virucide; hepatotropic; antiinflammatory; blood plasma processing; KKNS4a.
                                                                Gaps
                                                                ö
Query Match 100.0%; Score 49; DB 22; Length 23; Best Local Similarity 100.0%; Pred. No. 0.055; Matches 11; Conservative 0; Mismatches 0; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Peptide KKNS4a for HCV NS3 protease kinetic assay.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       (BRIM ) BRISTOL-MYERS SQUIBB PHARMA CO.
                                                                                                                                                                                                                                                                                                                                              ABG32508 standard; peptide; 23 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             13-DEC-2000; 2000US-255290P
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   12-DEC-2001; 2001WO-US47911
                                                                                                                                                                                                                                                                                                                                                                                                                                                                 15-NOV-2002 (first entry)
                                                                                                                             1 GSVVIVGRIVL 11
                                                                                                                                                                3 GSVVIVGRIVL 13
```

Gaps

```
The invention relates to pyrimidinones of a formula given in the claims of the specification, their stereoisomers, salts and prodrugs. In assays, the pyrimidinone compounds inhibited Hepatitis C virus (HCV) NS3 protease with IC_5 0 values of less than 100 micro M. The compounds are useful for treating viral infection e.g. HCV infection (the causative agent of acute hepatitis and associated with cirrhosis and liver cancer) and as a reagent used as inhibitors of HCV protease in the processing of blood plasma for diagnostic and other commercial purposes. The present sequence is a peptide, KNNS4a, used in an NS3 kinetic assay.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         salts or products which are useful as serine protease inhibitors.
The imidazolidinones of the invention may have virucide, hepatotropic, or antiinflammatory activities and may be used as a serine protease inhibitor (preferably Hepatitis C virus (HCV) NS3 protease inhibitor) or a HCV growth inhibitor. Compounds of the invention are useful for treating viral infection e.g. hepatitis C virus (HCV) infection and as a reagent used as inhibitors of HCV protease in the processing of blood plasma for diagnostic and other commercial purposes. The invention inhibit HCV NS3 protease and/or HCV HCV HCV NS3 protease and/or HCV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          This invention relates to novel imidazolidinones or their stereoisomers,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             growth and thus can be used in the blood plasma assay. The present sequence represents the KKNS4a peptide used in enzyme assay experiments in the examples of the specification.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          New imidazolidinones useful as serine protease inhibitors in the treatment of \mathbf{e}.\mathbf{g}. viral infection
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           HCV; hepatitis C; imidazolidinone; serine protease inhibitor;
virucide; hepatotropic; antiinflammatory; NS3 protease; KKNS4a;
growth inhibitor; viral infection; blood plasma processing.
                                                                                                                                                                                                                                                                                                                     ö
                                                                                                                                                                                                                                                                           100.0%; Score 49; DB 23; Length 23; 100.0%; Pred. No. 0.055; ive 0; Mismatches 0; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           (BRIM ) BRISTOL MYERS SQUIBB PHARMA CO.
Example 140; Page 192; 270pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Example 20; Page 112; 173pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ABG31914 standard; Peptide; 23 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                12-DEC-2001; 2001WO-US47916
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      13-DEC-2000; 2000US-255168P
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                (first entry)
                                                                                                                                                                                                                                                                                                  Local Similarity 100.
                                                                                                                                                                                                                                                                                                                                                           1 GSVVIVGRIVE 11
                                                                                                                                                                                                                                                                                                                                                                                  WPI; 2002-599498/64
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Ž
                                                                                                                                                                                                                                        23 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        KKNS4a peptide.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   W0200248157-A2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                23
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                05-NOV-2002
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        20-JUN-2002.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Synthetic
                                                                                                                                                                                                                                            Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ABG31914;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Sequence
                                                                                                                                                                                                                                                                               Query Match
                                                                                                                                                                                                                                                                                                      Best Loca
Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                             RESULT 10
                                                                                                                                                                                                                                                                                                                                                                                                                                                                              줨
                                                                                                                                                                                                                                                                                                                                                               ö
```

```
ö
                              ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    The present invention relates to hepatitis C virus (HCV) core antigen and NS (nonstructural) 3/4a antibody combination assay that can detect both HCV antigens and antibodies present in a sample using a single solid matrix as well as immunoassay solid supports for use in the assay. The solid support is useful for detecting HCV infection in a biological sample. The present sequence is NS4A peptide which is used to purify NS3/4a conformational epitope in the exemplification of the invention.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Gaps
                              Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ä
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Cost
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Immunoassay solid support useful for detecting hepatitis C virus infection in a biological sample, comprises at least one of HCV anti-core antibody and HCV NS3/4a epitope, bound to the support
                                                                                                                                                                                                                                                               Hepatitis C virus; NS3/4a antibody; HCV infection; NS4A peptide.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ;
0
                            ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Length 23;
 Length 23;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      George-Nasciemento C,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 100.0%; Score 49; DB 23; Length 23
100.0%; Pred. No. 0.055;
. . wismatches 0; Indels
                                                                                                                                                                                                                                     NS4A peptide used to purify NS3/4a conformational epitope.
                              Indels
100.0%; Score 49; DB 23;
100.0%; Pred. No. 0.055;
ive 0: Mismatches 0;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Tandeske L,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Example 2; Page 50; 87pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                AAU76376 standard; Peptide; 23 AA.
                                                                                                                                                      AAE18687 standard; peptide; 23 AA.
                                                                                                                                                                                                                                                                                                                                                                                                    15-JUN-2000; 2000US-212082P.
02-APR-2001; 2001US-280811P.
02-APR-2001; 2001US-280867P.
                                                                                                                                                                                                                                                                                                                                                                           14-JUN-2001; 2001WO-US19369.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       (first entry)
                                                                                                                                                                                                          (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Ouery Match
Best Local Similarity 100.
               Best Local Similarity 100.
Matches 11; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1 GSVVIVGRIVL 11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               3 GSVVIVGRIVL 13
                                                      1 GSVVIVGRIVL 11
                                                                                  3 GSVVIVGRIVL 13
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Arcangel P,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               WPI; 2002-179522/23
                                                                                                                                                                                                                                                                                                                                                                                                                                                            (CHIR ) CHIRON CORP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                23 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Medina-Selby A;
                                                                                                                                                                                                                                                                                                                    W0200196875-A2.
                                                                                                                                                                                                                                                                                          Unident1fled.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      08-MAY-2002
                                                                                                                                                                                                          17-MAY-2002
                                                                                                                                                                                                                                                                                                                                                 20-DEC-2001.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Seguence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         AAU76376;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Chien DY,
                                                                                                                                                                                AAE18687;
    Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     RESULT 12
                                                                                                                            RESULT 11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   AAU76376
                                                                                                                                          AAE18687
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  8×4×5
                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                               a
```

9

```
Best Local Similarity 100.
Matches 11; Conservative
                                                                                                                                   Hepatitis C virus type 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1 GSWIVGRIVL 11
                                                                                                                                                                                                                                                                                                                                                                                                                                                           Chien DY, Arcangel P, Medina-Selby A;
                                                                                                                                                                                                                                                                                                                                                                                                                               MEDINA-SELBY A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               GSVVIVGRIVL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              WPI; 2003-147573/14.
                                                                                                                                                                                                                                                                                                                                        CHIEN D Y.
ARCANGEL P.
TANDESKE L.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         23 AA;
                                                                                                                                                                    US2002146685-A1.
                                                                                                                                                                                                      10-OCT-2002
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        AAB15212;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Seguence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Query Match
                                                                                                                                                                                                                                                                                                                                                                            (TAND/)
(GEOR/)
(COIT/)
                                                                                                                                                                                                                                                                                                                                          CHIE/)
                                                                                                                                                                                                                                                                                                                                                                                                                               (MEDI/)
                                                                                                                                                                                                                                                                                                                                                          ARCA/)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   RESULT 14
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   AAB15212
ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     4×2×4×
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         The present invention relates to a new immunoassay solid support consisting essentially of at least one hepatitis C virus (HCV) NS3/4a conformational epitope and a multiple epitope fusion antigen (MEFA).

Bound to the support. The NS3/4a conformational epitope and/or sample from an HCV-Infected individual. The immunoassay of the invention is useful for detecting hepatitis C virus infection in a biological sample. The method of the invention provides a sensitive, accurate diagnostic and prognostic tool to provide adequate patient care and to prevent transmission of HCV by blood and by blood products, or by personal contact. Use of NS3/4a conformational epitope in combination with MEFA, provides a sensitive and reliable method for detecting early HCV serconversion. Use of MEFA has the aded advantages of decreasing masking problems, improving sensitivity in detecting antibodies by allowing a greater number of epitopes on a unit surface area of the incidence of false results is reduced because of the identification and the use of highly immunogenic HCV antigens which are present during the early stages of HCV serconversion. The present amino acid sequence represents the non-structural protein NS4A peptide sequence. The peptide
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ö
                                           Hepatitis C virus: HCV; NSJ/4a conformational epitope: seroconversion; immunoassay solid support: multiple epitope fusion antigen; MEFA; non-structural protein; NS4A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                             Immunoassay solid support, useful for detecting hepatitis C virus infection in biological sample, comprises HCV NS3/4a conformational epitope and multiple epitope fusion antigen bound to the support -
                                                                                                                                                                                                                                                                                                                                                            Coit D;
              Hepatitis C virus (non-structural protein) NS4A peptide sequence.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ö
                                                                                                                                                                                                                                                                                                                                                            Tandeske L, George-nasciemento C,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          100.0%; Score 49; DB 23; Length 23;
100.0%; Pred. No. 0.055;
.ive 0; Mismatches 0; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ABG72264 standard; peptide; 23 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Example 3; Page 48; 92pp; English.
                                                                                                                                                                                                                                                     15-JUN-2000; 2000US-212082P.
02-APR-2001; 2001US-280811P.
02-APR-2001; 2001US-280867P.
                                                                                                                                                                                                                      14-JUN-2001; 2001WO-US19156
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Dest Local Similarity 100.00
Best Local Similarity 100.00
Dest Local 11; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        06-MAR-2003 (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1 GSVVIVGRIVL 11
                                                                                                                                                                                                                                                                                                                                                          Arcangel P,
                                                                                                                                                                                                                                                                                                                                                                                                              WPI; 2002-090228/12.
                                                                                                                                                                                                                                                                                                                          (CHIR ) CHIRON CORP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           23 AA;
                                                                                                                   Hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                            Medina-selby A;
                                                                                                                                                     WO200196870-A2.
                                                                                                                                                                                      20-DEC-2001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ABG72264;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Seguence
                                                                                                                                                                                                                                                                                                                                                            Chien DY,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             epitope
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 RESULT 13
ABG72264
```

```
The present invention relates to immunoassays comprising Hepatitis C Virus (HCV) N33/4a conformational epitope and multiple epitope fusion antigen (MEFA), bound to a solid support. The N33/4a epitope fusion the multiple epitope fusion antigen react with anti-HCV antibodies present in a biological sample from an HCV-infected individual. The immunoassays and methods of the invention are useful for detecting support provides a sensitive and reliable method for detecting early HCV seroconversion. The assays can detect HCV infection caused by any six known genotypes of HCV. The use of the multiple epitope fusion proteins decreases masking problems, improves sensitivity in detecting
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Immunoassay solid support for detecting Hepatitis C Virus infection in biological samples, comprises Hepatitis C Virus conformational epitope and multiple epitope fusion antigen -
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               antibodies by allowing a greater number of epitopes on a unit area of substrate, and improves selectivity. The present sequence representing HCV type 1 (HCV-1) NS4a peptide is used in a protease enzyme activity assay in the examples of the present invention.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Coit D;
                                                                              Immunoassay solid support; Hepatitis C Virus type-1; HCV-1; MS3/4a conformational epitope; multiple epitope fusion antigen; MEFA; anti-HCV antibody; MS3/4a conformational antigen; HCV infection; E2 hypervariable region.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ;
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          100.0%; Score 49; DB 24; Length 23; 100.0%; Pred. No. 0.055; ive 0; Mismatches 0; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                George-Nascimento C,
Hepatitis C Virus type-1 (HCV-1) NS4a peptide.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Tandeske L,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               AAB15212 standard; protein; 195 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Example 3; Page 17; 45pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               GEORGE-NASCIMENTO C.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                15-JUN-2000; 2000US-212082P. 02-APR-2001; 2001US-280811P. 02-APR-2001; 2001US-280867P.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 14-JUN-2001; 2001US-0881654
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    19-DEC-2000 (first entry)
```

ö

ò a

```
ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                 The present sequence is a fusion protein created using the Hepatitis C virus (HCV) NS3 and NS4A protease enzymes. These proteins are both essential for the replication of the virus, acting to cleave its replicative proteins from the polyprotein produced from the HCV genome. Inhibitors of the two proteins should be effective as antiviral treatments of HCV infection. This is useful as HCV can lead to chronic liver disease such as cirrhosis, liver failure and liver cancer. The present invention concerns a number of MS3 metants and MS3-MS4 fusion proteins which can be used to identify inhibitors of this type, as well as enabling structural studies of the protease and protease:inhibitor
                                                                                                                                                                                                                                                                                                                                              Modified hepatitis C virus (HCV) NS3 protease comprising at least 1 substitution of a hydrophobic alpha-helix 0 amino acid to a hydrophilic amino acid, useful for screening inhibitors that may treat hepatitis C
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Hepatitis; NS3 protease; viral replication; chronic liver disease;
liver failure; liver cancer; mutant; mutein.
                         Hepatitis; NS3 protease; viral replication; chronic liver disease;
liver failure; liver cancer.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ó
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          100.0%; Score 49; DB 21; Length 195; 100.0%; Pred. No. 0.56; ive 0; Mismatches 0; Indels 0
                                                                                                                                                                                                                                                                        Coldfarb V;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Hepatitis C virus NS4A-NS3 fusion protease #2.
Hepatitis C virus NS4A-NS3 fusion protease #1.
                                                                                                                                                                                                                                                                      Zhang Y,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               AAB15220 standard; protein; 195 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                        Example 2; Fig 10; 66pp; English.
                                                                                                                                                                                                                                          (BRIM ) BRISTOL-MYERS SQUIBB CO
                                                                                                                                                                                                                                                                        Wittekind M, Weinheimer S,
                                                                                                                                                                                                            99US-0115271
                                                                                                                                                                                06-JAN-2000; 2000WO-US00345
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       19-DEC-2000 (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1 GSVVIVGRIVL 11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    GSVVIVGRIVL 15
                                                                                                                                                                                                                                                                                                    2000-465976/40.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Query Match
Best Local Similarity
Matches 11; Conser
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                195 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Hepatitis C virus Synthetic.
                                                                           Hepatitis C virus
Synthetic.
                                                                                                                                                                                                                                                                                                                    N-PSDB; AAA73328
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             WO200040707-A1
                                                                                                                      WO200040707-A1
                                                                                                                                                                                                            08-JAN-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         13-JUL-2000.
                                                                                                                                                  13-JUL-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     complexes
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           AAB15220
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 RESULT 15
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     셤
```

```
ö
                                                                                                                                                                                                                                                                                                                            The present sequence is a mutated version of a fusion protein created using the Hepatitis C virus (HVV) NS3 and NS4A protease enzymes. These proteins are both essential for the replication of the virus, acting to cleave its replicative proteins from the polyprotein produced from the HCV genome. Inhibitors of the two proteins should be effective as antiviral treatments of the two proteins should be effective as antiviral treatments of HCV infection. This is useful as HCV can lead to chronic liver disease such as cirrhosis, liver failure and liver cancer. The present invention concerns a number of NS3 mutants and NS3-NS4A well as enabling structural studies of the protease and protease:inhibitor complexes. This sequence contains the alpha-helix0-1
                                                                                                                                                                                                  Modified hepatitis C virus (HCV) NS3 protease comprising at least 1 substitution of a hydrophobic alpha-helix 0 amino acid to a hydrophilic amino acid, useful for screening inhibitors that may treat hepatitis C
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        100.0%; Score 49; DB 21; Length 195; 100.0%; Pred. No. 0.56; tive 0; Mismatches 0; Indels
                                                                                                            Goldfarb V;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            completed: August 30, 2003, 19:12:26
                                                                                                            Zhang Y,
                                                                                                                                                                                                                                                                                               Claim 23; Fig 12; 66pp; English.
                                                                       (BRIM ) BRISTOL-MYERS SQUIBB CO
                                                                                                            Wittekind M, Weinheimer S,
06-JAN-2000; 2000WO-US00345
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   11; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1 GSVVIVGRIVL 11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         WPI; 2000-465976/40.
N-PSDB; AAA73329.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           195 AA;
                                  08-JAN-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Search co
a
```

```
GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.
```

- protein search, using sw model OM protein August 30, 2003, 19:02:22; Search time 0.905317 Seconds Run on:

(Without alignments) 1168.492 Million cell updates/sec

US-09-965-594-26 49 1 GSVVIVGRIVL 11 Perfect score: Sequence: BLOSUM62 Gapop 10.0 , Gapext 0.5 Scoring table:

283308 segs, 96168682 residues Searched:

Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0 Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0% Maximum Match 100% Listing first 45 summaries

PIR\_76:\* Database

1: pirl:# 2: pirl:# 3: pir2:# 4: pir4:#

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

æ

Description	genome polyprotein			genome polyprotein	genome polyprotein	0	•	genome polyprotein			genome polyprotein		probable copper-tr	probable dna polym	valine-tRNA ligase	H-2 class I histoc	. probable membrane	env polyprotein -	hypothetical prote	hypothetical prote	Ψ	env polyprotein pr	env polyprotein pr					rotein	Ü
ΩI	GNWVTC	GNWVCJ	A45573	S18030	GNWVTW	PS0326	PC2219	GNWVCH	S40770	JC5620	GNWVC3	J01366	G7097B	T41395	E89949	PS0117	S54580	C46357	E71234	T05630	VCLJEV	VCLJ22	VCLUEW	VCLJE1	VCLJE2	VCLJE3	VCLJWS	VCLJE4	PS0130
DB	-	-	7	-	Н	a	7	-	٦	7	-	ď	7	7	N	~	ď	7	7	7	-	~	~	-	~	٦	-	-	~
Length	3010	3010	3010	3010	3010	492	876	3011	3011	3014	3011	716	718	574	876	44	196	N	570	4	2	S	2	S	S	5	859	9	37
Ouery Match	98.0	98.0	98.0	0.86		83.8	89.8	868.8		89.8	87.8	85.7	75.5	71.4	71.4	<b>9</b> .69	4.69	69.4	<b>4</b> .69	69.4	<b>9</b> .69	69.4	69.4	4.69	69.4	69.4	69.4	69.4	67.3
Score	8	48	48	48	48	44	44	44	4	4	43	42	37	35	35	34	34	34	34	34	34	34	34	34	34	34	34	34	33
Result No.		7	3	₹	Ś	9	7	<b>∞</b>	6	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	92	27	28	53

H-2 class I histoc H-2 class I histoc	H-2 class I-like h genome polyprotein	genome polyprotein conserved hypothet	probable tail asse 3-oxoacvl-(acvl-ca	conserved domain p	hypothetical prote methenyltetrahydro	lipoprotein [impor hypothetical prote	hypothetical prote argininosuccinate
PS0127 PS0118	A32934 PC1307	PC1306 D75611	D90908 AG3213	F95091 B97959	C85631 F69491	F89771 B71407	E90047 C83204
0.0	~ ~	7 7	77	77	~ ~	9	77
37	39	209	226 248	271 291	300	331 386	395 405
67.3	67.3	67.3	67.3	67.3 67.3	67.3 67.3	67.3 67.3	67.3
333	33	333	333	333	333	33	33
30 31	32 33	3.4 5.4	36 37	38 39	<b>4</b> 0	<b>4.4</b> 3.2	44 45

### ALIGNMENTS

```
genome polyprotein - hepatitis C virus
N:Contains: capsid protein C; envelope protein M; hepacivirin (EC 3.4.21.98) (nonstranctural protein NS4s; nonstructural protein NS4s; nonstructural protein NS4s; nonstructural protein SC; Species: hepatitis C virus
C;Species: hepatitis C virus
C;Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 19-Jan-2001
                                                                                                                                                                                                                                                                                                                                                                       ö
                                                                                                                                                                                                                                                                                                                                                                        Gaps
                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                               Score 48; DB 1; Length 3010;
Pred. No. 2.4;
                                                                                                                                                                                                                                                                                                                                                                       0; Indels
                                                                                                                                                                                                                                                                                                                                                                       1; Mismatches
                                                                                                                                                                                                                                                                                                                                              98.08;
90.9<del>8</del>;
                                                                                                                                                                                                                                                                                                                                                                                                          Best Local Similarity 90.9 Matches 10; Conservative
                                                                                                                                                                                                                                                                                                                                                                                            1 GSVVIVGRIVL 11
                                                                                                                                                                                                                                                                                                                                                 Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                 q
                                                                                                                                                                                                                                                                                                                                                                                               ò
```

General polyprotein - hepatitis C virus (strain J)
N:Contains: capsid protein C; envelope protein M; major envelope protein E; nonstructorians: capsid protein C; envelope protein NS4b; nonstructural protein NS5
C;Species: hepatitis C virus
C;Species: hepatitis C virus
C;Accession A3925; ps0086
C;Accession A3925; ps0086
C;Accession M; Hijikata, M; Ootsuyama, Y; Nakagawa, M; Ohkoshi, S; Sugimura, T; S}
Proc. Natl: Acad. Sci. U.S.A. 87, 9524-9528, 1990
A;Title: Molecular cloning of the human hepatitis C virus genome from Japanese patit

ă,

```
A. Molecule type: genomic RNA
A. Residues: 1.547, T'.549-621, V', 623-624, S', 626-652, DL', 655-761, T', 763-782 (HOW)
A. Residues: 1.547, T'.549-621, V', 623-624, S', 626-652, DL', 655-761, T', 763-782 (HOW)
A. Cross-references: EMBL.X61591
A. Note: this sequence is inconsistent with the nucleotide translation
A. Note: the authors translated the codon AGG for residue 43 as Pro, TGG for residue:
A. Note: the authors translated the codon AGG for residue 43 as Pro, TGG for residue:
A. Note: sequence extracted from NOBI backbone (NOBIN:121747, NCBIP:121748)
C. Superfamily: hepatitis C virus genome polyprotein
C. Reywords: ATP: 91ycoprotein: Mydrolass: nucleotide binding: P-loop; polyprotein: F: 115.7Product: capsid protein C *status predicted CRD>
F: 116-191/Product: capsid protein C *status predicted CRD>
F: 192-389/Product: major envelope protein E *status predicted CNS>
F: 300-729/Product: monstructural protein NSI *status predicted CNS>
F: 1007-1615/Product: hepacivirin *status predicted CNS>
F: 1230-1237/Region: nucleotide binding motif B (P-loop)
F: 7312-1017/Region: nucleotide binding motif B
F: 7312-01217/Region: nucleotide binding motif B
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        genome polyprotein – hepatitis C virus (strain Talwan)
N:Contains: capsid protein C; envelope protein M; hepacivirin (EC 3.4.21.98) (nonstru
protein NS4a; nonstructural protein NS4b; nonstructural protein NS5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 genome polyprotein - hepatitis C virus (isolate JK1)
N;Contains: capsid protein C; envelope protein M; hepacivirin (EC 3.4.21.98) (nonstri
protein NS4a; nonstructural protein NS4b; nonstructural protein NS5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           A;Experimental source: isolate JK1 from an individual
R;Honda, M.; Kaneko, S.; Unoura, M.; Kobayashi, K.; Murakami, S.
Arch. Virol. 128, 163-169, 1993
A;Title: Sequence analysis of putative structural regions of hepatitis C virus isola'
A;Reference number: A48332; MUID:93119270; PMID:8380322
A;Accession: S33570
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            F:1616-1862/Product: nonstructural protein NS4a #status predicted <N4A>
F:1863-2011/Product: nonstructural protein NS4 #status predicted <N4B>
F:2014-3010/Product: nonstructural protein NS5 #status predicted <NS5>
F:196,209,234,250,305,417,423,448,532,540,556,576,623,645/Binding site: carbohydrate
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Cispecies: hepatitis C virus
A; Variety: isolate JK1
C; Date: 19-May-2000 #sequence_revision 19-May-2000 #text_change 23-Mar-2001
C; Date: 19-May-2000 #sequence_revision 19-May-2000 #text_change 23-Mar-2001
C; Date: 19-May-2000 #sequence_revision 18029
R; Honda, M.; Kaneko, S.; Masashi, U.; Kobayashi, K.; Murakami, S.
submitted to the BMBL Data Library, September 1991
A; Description: A whole genome of hepatitis C virus CDNA was isolated from a single |
A; Reference number: S18028
A; Accession: S18029
                                                                                                                                    ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           C;Species: hepatitis C virus
A;Note: host Homo sapiens (man)
C;Date: 31-Dec-1992 *sequence_revision 31-Dec-1992 *text_change 19-Jan-2001
                                                                                                                                    Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        A;Molecule type: genomic RNA
A;Residues: 1-3010 <HON>
A;Cross-references: EMBL:X61596; NID:959478; PIDN:CAA43793.1; PID:959479
                                                                                                                                    ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Score 48; DB 1; Length 3010;
            Length 3010;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Indels
                                                                                                                                    Indels
            DB 1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Pred. No. 2.4;
1; Mismatches
            Score 48; DB Pred. No. 2.4;
                                                                                                                                    1; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             98.08;
90.98;
            98.08;
                                                                                                                                                                                                                                                                                                             1678 GSVVIVGRIIL 1688
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1678 GSVVIVGRIIL 1688
Query Match
Best Local Similarity 90.9
Matches 10; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             F;1316-1319/Region: DEXH motif
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Best Local Similarity 90.9
Matches 10; Conservative
                                                                                                                                                                                                                                              1 GSVVIVGRIVL 11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1 GSVVIVGRIVL 11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    C; Accession: A40244
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       RESULT 5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            GNWVTW
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     A. Accession: A4573
A. Status: preliminary
A. McCession: A4573
A. Status: preliminary
A. McCoss. Preliminary
A. McCoss. Preliminary
A. McCoss. Presences: GB:D01171; NID:g221612; PIDN:BAA01943.1; PID:g221613
A. Residues: 1-3010 cTANY
A. Residues: 1-3010 cTANY
A. Residues: 1-3010 cTANY
A. McCoss. Feferences: GB:D011168; GB:D01171; NID:g221612; PIDN:BAA01943.1; PID:g221613
A. Reperimental source: HCV-TT
A. Mote: sequence extracted from NCBI backbone (NCBIN:106206, NCBIP:106207)
C. Superimental source: HCV-TT
A. Mote: sequence extracted from NCBI backbone (NCBIN:106206, NCBIP:106207)
C. Superimental source: Richard Setatus predicted cCPC>
F:116-191/Product: envelope protein M *status predicted cMEE>
F:192-389/Product: nonstructural protein NSI *status predicted cNSI>
F:1007-1615/Product: nonstructural protein NSI *status predicted cNSI>
F:1007-1615/Product: nonstructural protein NSI *status predicted cNSI>
F:1310-1137/Region: nucleotide-binding motif A (P-loop)
F:1312-1317/Region: nucleotide-binding motif B
F:1312-1317/Region: nucleotide-binding motif B
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                A.Accession: PS0086
A.Molecule type: genomic RNA
C.Comment: The cleavage sites of this polyprotein bave not been determined.
C.Comment: The cleavage sites of this polyprotein
C.Comment: The cleavage sites of this polyprotein polyprotein
C.Comment: The cleavage sites of this protein the status predicted constant and protein NSI *status predicted constant and protein NSI *status predicted constant and protein NSI *status predicted constant and const
                                            A;Accession: A39253
A;Molecule type: genomic RNA
A;Molecule type: genomic RNA
A;Molecule type: genomic RNA
A;Molecule type: 1-3010 < KAI>-
A;Molecule type: 1-3010 < KAI>-
A;Molecule type: 1-3010 < KAI>-
A;Cross-references: GB:D90208; NID:9221610; PIDN:BAA14233.1; PID:9221611
R;Kato, N.; Ohkoshi, S.; Shimotohno, K.
R;Kato, N.; Ohkoshi, S.; Shimotohno, K.
A;Title: Japanese 159021423, 1989
A;Title: Japanese 1501ates of the non-A, non-B hepatitis viral genome show sequence variance mumber: PS0085
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     genome polyprotein - hepatitis C virus (strain JT)
N;Contains: capsid protein C: envelope protein M; hepacivirin (EC 3.4.21.98) (nonstructuor NS4 strong NS5 strong NS4 strong NS4 strong NS4 strong NS4 strong NS4 strong NS5 strong NS4 strong NS4 strong NS5 strong NS6 strong N
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            F;1616-1862/Product: nonstructural protein NS4a #status predicted <N4A>
F;1863-2013/Product: nonstructural protein NS4b *status predicted <N4B>
F:2014-3010/Product: nonstructural protein NS5 *status predicted <NS5>
F:196,209,234,250,305,325,417,423,430,448,532,556,576,623,645,1213,1255,2041,2077,2240,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           F;1616-1862/Product: nonstructural protein NS4a #status predicted <N4A>F;1863-2013/Product: nonstructural protein NS4b #status predicted <N4B>F;2014-3010/Product: nonstructural protein NS5 #status predicted <NS5>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Score 48; DB 1; Length 3010;
Pred. No. 2.4;
1; Mismatches 0; Indels
A; Reference number: A39253; MUID: 91088550; PMID: 2175903
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             98.08;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         F;1316-1319/Region: DEXH motif
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Query Match
Best Local Similarity 90.9
Matches 10; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1678 GSVVIVGRIII, 1688
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1 GSVVIVGRIVL 11
```

q

Gaps

ö

Length 876; Indels

```
F:192-247/Product: E1 (carboxyl end) *status predicted <ERE>
F:248-411/Product: E2/NS1 (amino end) *status predicted <ENR>
F:248-318/Region: E2/NS1 (amino end)
F:339-411/Region: NS1 *status predicted <NSR>
F:412-783/Product: NS3 *status predicted <NSR>
F:788-783/Product: NS4 *status predicted <NSR>
F:788-876/Product: NS4 *status predicted <NSR>
F:781.278/Product: NS4 *status predicted <NSR>
F:781.287/Product: NS4 *status predicted <NSR>
F:781.287,294,312,340/Binding site: carbohydrate (Asn) (covalent) *status predicted
                                                                   A;Cross-references: GB:L29577; GB:L29578; GB:L29579
A;Experimental source: serum
C;Superfamily: hepatitis C virus genome polyprotein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Score 44; DB 2;
Pred. No. 4;
1; Mismatches
                                                                                                                                                                      C;Reywords: glycoprotein
F;1-191/Product: core #status predicted <COE>
F;68-78/Region: variable
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      89.8%;
81.8%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          89.8%;
90.9%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         F;1316-1319/Region: DEXH motif
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1678 GCWIVGRIVL 1688
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Query Match
Best Local Similarity 81.8
Matches 9; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         C; Species: hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1 GSWINGRIVL 11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               A; Reference number: A36814
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Query Match
Best Local Similarity
Matches 10; Conserv
          A; Molecule type: mRNA
A; Residues: 1-876 <STU>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 A; Accession: A36814
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               N; Contains
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   genome
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 GNWVCH
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                   A; Accession: A40244
A; Molecule type: genomic RNA
A; Recidues: 13010 < CHED
A; Recidues: 13010 < CHED
A; Residues: 13010 < CHED
A; Superfamily: hepatitis C virus genome polyprotein
C; Superfamily: hepatitis C virus genome polyprotein; glycoprotein; hydrolase; nonstructura
C; Reyords: ATP: capsid protein C *status predicted < CPD
F; 116-131/Product: envelope protein B *status predicted < CPD
F; 116-131/Product: major envelope protein B$ *status predicted < NES>
F; 130-1006/Product: nonstructural protein NS1 *status predicted < NS2>
F; 1307-1615/Product: nonstructural protein NS2 *status predicted < NS2>
F; 1312-1317/Region: nucleotide-binding motif B
F; 1316-1319/Region: nucleotide-binding motif B
F; 1316-1319/Region: nucleotide-binding motif B
F; 1316-1319/Region: nonstructural protein NS4 *status predicted < NAA>
F; 1663-2013/Product: nonstructural protein NS5 *status predicted < NAS>
F; 1014-3010/Product: nonstructural protein NS5 *status predicted < NS5>
F; 196, 209, 233, 234, 250, 305, 325, 417, 423, 430, 448, 532, 540, 556, 576, 623, 645, 1213, 1207
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  C.Accession: PS0326
Fili, U.S. Tong, S.P.; Vitvitski, L.; Lepot, D.; Trepo, C.
Gene 105, 167-172, 1991
A.Title: Two French genotypes of hepatitis C virus: homology of the predominant genotype
A.Reference number: PS0326; MUID:92039028; PMID:1718820
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         polypeptide - hepatitis C virus (type 5a) (fragments)
N:Contains: core protein; El (carboxyl end); E2/NS1 (amino end); NS3 protein; NS4A prote
C;Species: hepatitis C virus
C;Date: 20-Feb-1995 *sequence_revision 20-Feb-1995 *text_change 17-Nov-2000
C;Accession: PC2219
R;Stuyver, L.; Arnhem, W.V.; Wyseur, A.; Maertens, G.
R;Chen, P.J.; Lin, M.H.; Tai, K.F.; Liu, P.C.; Lin, C.J.; Chen, D.S. Virology 188, 102-113, 1992
A;Title: The Taiwanese hepatitis C virus genome: sequence determination and mapping the A;Title: The Taiwanese hepatitis C virus genome: sequence determination and mapping the A;Reference number: A40244; MUID:92230206; PMID:1314449
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              R;Stuyver, L.; Arnhem, W.V.; Wyseur, A.; Maertens, G.
Blochem. Biophys. Res. Commun. 202, 1308-1314, 1994
A;Title: Cloning and phylogenetic analysis of the core, E2, and NS3/NS4 regions of the
A;Reference number: PC2219; MUID:94338342; PMID:7520237
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      polyprotein - hepatitis C virus (isolate Fla) (fragments)
C:Species: hepatitis C virus
C:Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 17-Nov-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            0;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Score 48; DB 1; Length 3010;
Pred. No. 2.4;
1; Mismatches 0; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       A; Accession: PS0326
A; Molecule type: genomic RNA
A; Residues: 1-492 <LID3
A; Cross-references: GB: M60220
A; Note: this sequence corresponds to nonstructural protein NS3:
A; Note: translation of the nucleotide sequence is not complete
C; Superfamily: hepatitis C virus genome polyprotein
C; Keywords: polyprotein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Score 44; DB 2; Length 492;
Pred. No. 2.3;
0; Mismatches 1; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  89.8%;
90.9%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            98.08;
90.98;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Guery Match
Best Local Similarity 90.9'
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1678 GSVVIVGRIIL 1688
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Query Match 89.8
Best Local Similarity 90.9
Matches 10; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     229 GCWIVGRIVI. 239
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1 GSVVIVGRIVL 11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1 GSVVIVGRIVL 11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             õ
```

```
A; Note: host Homo sapiens (man)
C;Date: 31-Dec-1992 #sequence_revision 31-Dec-1992 #text_change 19-Jan-2001
C;Accession: A36814; A41546
R;Inchauspe, G:: Zebedee, S.; Lee, D.H.; Sugltani, M.; Nasoff, M.; Prince, A.M.
S;Inchauspe, G: Zebedee, S.; Lee, D.H.; Sugltani, M.; Nasoff, M.; Prince, A.M.
Submitted to GenBank, July 1992
A;Description: Genomic structure of the human prototype strain H of hepatlis C viru
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Accordance: annotation
Associated as a consistency of the consistency 
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           A; Molecule type: genomic RNA
A; Residues: 1-3011 - TINC>
A; Cross-references: GB: MG7463; NID: 9329737; PIDN: AAA45534.1; PID: 9329738
A; Cross-references: GB: MG7463; NID: 9329737; PIDN: AAA45534.1; PID: 9329738
B; Inchauspe, G.; Zebedee, S.; Lee, D.H.; Sugitani, M.; Nasoff, M.; Prince, A.M.
Proco: Natl. Acad. Sci. US.As. 88, 10292-10296, 1991
Proco: Natl. Acad. Sci. US.As. 88, 10292-10296, 1991
A; Title: Genomic structure of the human prototype strain H of hepatitis C virus: com
A; Reference number: A41546; MUID: 92052256; PMID: 1658800
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  F;1616-1862/Product: nonstructural protein NS4a #status predicted <N4A>
F;1863-2013/Product: nonstructural protein NS4b #status predicted <N4B>
F;2014-3011/Product: nonstructural protein NS5 #status predicted <NS5>
F;2034-3011/Product: nonstructural protein NS5 #status predicted <NS5>
F;196,209,234,305,325,417,423,430,448,476,532,540,556,576,623,645,1213,1255,2041,224
jenome polyprotein - hepatitis C virus (strain H)
N;Contains: capsid protein C; envelope protein M; hepacivirin (EC 3.4.21.98) (nonstr
protein NS4a; nonstructural protein NS4b; nonstructural protein NS5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Score 44; DB 1; Length 3011;
Pred. No. 13;
0; Mismatches 1; Indels
```

Gaps

```
genome polyprotein - hepatitis C virus (strain HCV-1)
N'Contains: capsid protein C; envelope protein M; hepativitin (EC 3.4.21.98) (nonstru
N'Contains: capsid protein C; envelope protein M; hepativitin (EC 3.4.21.98) (nonstru
Drotein NS4s; nonstructural protein NS4b; nonstructural protein NS5
C; Species: hepatitis C virus
C; Daceles: hepatitis C virus
C; Accession: A39166; PQ0403; P00404
R; Choo, O.L.; Richman, K.H.; Han, J.H.; Berger, K.; Lee, C.; Dong, C.; Gallegos, C.;
Proc. Natl. Acad. Sci. U.S.A. 88, 2451-2455, 1991
A; Title: Genetic organization and diversity of the hepatitis C virus.
A; Reference number: A39166; MUID:91172826; PMID:1848704
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                A, Molecule type: mRNA
A, Residues: 1:3011 <CHO>
A, Cross-references: 08:M62321; NID:9329873; PIDN:AAA45676.1; PID:9329874
A, Cross-references: 08:M62321; NID:922
Gen. Virol. 73, 1131-1141, 1992
A, Title: Analysis of a new hepatitis C virus type and its phylogenetic relationship 'A, Reference number: PQ0393; MUID:92268871; PMID:136939
                                     F:731-1007/Product: nonstructural protein NSZ #status predicted <NS2>
F:1008-1616/Product: hepacivirin #status predicted <NS3>
F:1031-1338/Region: nucleotide-binding motif A (P-loop)
F:1317-1320/Region: nucleotide-binding motif B F:1317-1320/Region: nucleotide-binding motif B F:1317-1320/Region: nucleotide-binding motif F:183/Product: nonstructural protein NS4# #status predicted <N4A>
F:1864-2014/Product: nonstructural protein NS5 #status predicted <N4B>
F:2015-3014/Product: nonstructural protein NS5 #status predicted <NS5>
F:2210-2249/Region: interferon sensitivity determining #status predicted
F;390-730/Product: nonstructural protein NS1 *status predicted <NS1>F;731-1007/Product: nonstructural protein NS2 *status predicted <NS2>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              89.8%; Score 44; DB 1; Length 3014;
81.8%; Pred. No. 13;
Live 1; Mismatches 1; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               A: Experimental source: isolates E-bl6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       A; Molecule type: genomic RNA
A; Residues: 1577-1633 <CHA>
A; Cross·references: DDBJ;D10128
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            F,1316-1319/Region: DEXH motif
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1678 GCVVIVGRVVL 1688
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1679 GSVAIVGRIIL 1689
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Best Local Similarity 81.8
Matches 9; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Query Match 89.8
Best Local Similarity 81.8
Matches 9; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1 GSVVIVGRIVL 11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1 GSVVIVGRIVL 11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      A;Status: preliminary
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     A; Accession: A39166
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     A; Accession: PQ0404
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                RESULT 11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        đ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Accession: 30.20 Multiplys (25.44-49, 1997)

Biochem. Blophys Res. Commun. 236, 44-49, 1997

A:Reference number: JC5620; MUID:97366593; PMID:9223423

A:Accession: JC5620

A:Accession: JC562
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        A.Cross-references: GB:D00831; NID:g221511; PIDN:BAA00705.1; PID:g221512
A.Experimental source: isolate HC-J1
C.Superfamilty: hepatitis C virus genome polyprotein
F.2-115/Product: capsid protein C *status predicted <CPC>
F.3-115/Product: envelope protein M *status predicted <MED>
F.3-192-389/Product: major envelope protein E *status predicted <MED>
F.3-192-39/Product: nonstructural protein NS *status predicted <NS1>
F.3-10-106/Froduct: nonstructural protein NS *status predicted <NS2>
F.3-10-105/Froduct: hepacivirin *status predicted <NS2>
F.3-13-1337/Region: nucleotide-binding motif B
F.3-1312-1317/Region: nucleotide-binding motif B
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               genome polyprotein - hepatitis C virus (isolate EUH1480)
N;Contains: capaid protein C; envelope protein M; hepacivirin (EC 3.4.21.98) (nonstructu protein NS4a; nonstructural protein NS4b; nonstructural protein NS5 C; Species i hepatitis C virus
C; Date: 19-May-2000 *sequence_revision 19-May-2000 *text_change 19-Jan-2001
C; Accession: JC5620
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      A.Cross-references: EMEL:D10749; NID:9221586; PIDN:BAA01582.1; PID:9221587
R;Okamoto, H.; Okada, S.; Sugiyama, Y.; Yotsumoto, S.; Tanaka, T.; Yoshizawa, H.; Tsuda, Dho J. Exp. Med. 60, 167-177, 1990
A;Title: The 5'-terminal sequence of the hepatitis C virus genome.
A;Reference number: PC1284; MUID:91013116; PMID:2170712
A;Accession: PC1285
A;Aocession: PC1286
A;Aocession: PC1286
A;Residues: 1-513 < OK2>
                                                                                                                                                                                                              protein M; hepacivirin (EC 3.4.21.98) (nonstructu
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ö
                                                                                                                                                  genome polyprotein - hepatitis C virus
N;Contains: capsid protein C; envelope protein M; hepacivirin (EC 3.4.21.98
protein NG4s; nonstructural protein NS4b; nonstructural protein NS5
C;Species: hepatitis C virus
C;Date: 19-May-2000 #sequence_revision 19-May-2000 #text_change 19-Jan-2001
C;Accession: S40770; PC1285
R;Okamoto, H.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            F;1316-1319/Region: DEXH motif
F;1316-1319/Region: DEXH motif
F;1616-1862/Product: nonstructural protein NS4a #status predicted <N4A>
F;2014-3011/Product: nonstructural protein NS5 *status predicted <N85>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 89.8%; Score 44; DB 1; Length 3011; 90.9%; Pred. No. 13; 1; Indels 1; of Mismatches 1; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      F;384-408/Region: hypervariable #status predicted
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  submitted to the EMBL Data Library, March 1992
A;Reference number: $40770
A;Accession: $40770
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1678 GCVVIVGRIVL 1688
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Local Similarity 90.9
es 10; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                A; Molecule type: genomic RNA A; Residues: 1-3011 < OKA>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1 GSVVIVGRIVL 11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Best Loc
Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              RESULT 10
                                                           RESULT 9
```

```
A:Relegates: 1577-1633 ACR42.
A:Residues: 1577-1633 ACR42.
A:Resperimental source: isolates E-bl7
C:Superfamily: hepatitis C virus genome polyprotein
C:Superfamily: hepatitis C virus genome polyprotein; dlycoprotein; hydrolase; nonstruct
C:Keywords: ATP: capsid protein G vertains predicted cCPC.
F:115-Product: capsid protein M *status predicted cEPM.
F:106-191/Product: envelope protein M *status predicted cMED.
F:190-729/Product: nonstructural protein NSI *status predicted cMSIP.
F:300-729/Product: nonstructural protein NSI *status predicted cNSIP.
F:1070-1006/Product: nonstructural protein NSI *status predicted cNSIP.
F:1230-1237/Region: nucleotide-binding motif A (P-loop)
F:1312-1317/Region: nucleotide-binding motif A
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  F;1616-1862/Product: nonstructural protein NS4a *status predicted <N84>
F;1863-2013/Product: nonstructural protein NS4b *status predicted <NN8>
F;2014-3011/Product: nonstructural protein NS5 *status predicted <NNS>
F;2014-3011/Product: nonstructural protein NS5 *status predicted <NNS>
F;196,209,234,305,325,417,423,430,448,476,532,540,556,576,623,645,1213,1255,2041,2077
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          87.8%; Score 43; DB 1; Length 3011;
81.8%; Pred. No. 20;
iive 1; Mismatches 1; Indels
```

Gaps

;

Length 574; Indels

```
C; Species: Staphylococcus aureus
C;Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 22-Oct-2001
C;Accession: B89949
R;Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; (
ma, A.; Mizutani-Ui, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, F
C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.
Lancet 357, 1225-1240, 2001
A;Molecule type: DNA
A;Residues: 1-574 <MO>>
A;Cross-references: EMBL:AL023704; PIDN:CRA19261.1; GSPDB:GN00068; SPDB:SPCC553.09c
A;Experimental source: strain 972h-; cosmid c553
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-876 <KUR>
A;Cross-references: GB:BA000018; PID:g13701460; PIDN:BAB42754.1; GSPDB:GN00149
A;Experimental source: strain N315
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          A;Title: Whole genome sequencing of meticillin-resistant Stapylococcus aureus.
A;Reference number: A89758; MUID:21311952; PMID:11418146
A;Accession: E89949
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     valine-tRNA ligase [imported] - Staphylococcus aureus (strain N315)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         71.4%; Score 35; DB 2; Length 876; 70.0%; Pred. No. 1.8e+02; tive 2; Mismatches 1; Indels
                                                                                                                                                                                                                                                                                                                                                                              / Match 71.4%; Score 35; DB 2; 1
Local Similarity 77.8%; Pred. No. 1.2e+02;
hes 7; Conservative 2; Mismatches 0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Search completed: August 30, 2003, 19:20:34 Job time : 2.90532 secs
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         C; Superfamily: valine-tRNA ligase
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Best Local Similarity 70.0 Matches 7; Conservative
                                                                                                                                                                                                                                                                A;Map position: 3
A;Introns: 89/2; 415/3; 518/3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     2 SWIVGRIVL 11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      205 VVVGRIVV 213
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          3 VVIVGRIVL 11
                                                                                                                                                                                C;Genetics:
A;Gene: SPDB:SPCC553.09c
                                                                                                                                                                                                                                                                                                                                                                                       Query Match
Best Local
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 RESULT 15
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               probable dna polymerase alpha-primase associated subunit - fission yeast (Schizosacchard C;Species: Schizosaccharomyces pombe C;Species: 3chizosaccharomyces pombe C;bate: 03-bec-1999 *sequence_revision 03-bec-1999 *text_change 03-bec-1999 R;Wood, V; Rajandream. M.A. Harrell P.C.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      A:Molecule type: genomic RNA
A:Residues: 1.716 (KRE>
C:Superfamily: hepatitis c virus genome polyprotein
C:Keywords: glycoprotein: polyprotein
F:84,90,97,115,143,199,223,243,290,312/Binding site: carbohydrate (Asn) (covalent) #stat
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 probable copper-transporting atpase 11/9 - Mycobacterium tuberculosis (strain H37RV) (Species: Mycobacterium tuberculosis (Cipate: 17-ul-1998 #sequence_revision 17-Jul-1998 #text_change 18-Aug-2000 (Cipate: 17-ul-1998 #sequence_revision 17-Jul-1998 #text_change 18-Aug-2000 (Cipate: 17-Jul-1998 #text_change 18-Aug-200 (Cipate: 17-Jul-1998 #text_cha
                                                                                                                                                                                                                                                                                                   C;Accession: JQ1366
R;Kremsdorf, D.; Porchon, C.; Kim, J.P.; Reyes, G.R.; Brechot, C.
J Gen. Virol. 72, 2557-2561, 191
A;Title: A;Title: Total nucleotide sequence analysis of a French hepatitis C virus: implication A;Reference number: JQ1366; MUIO:92013977; PMID:1655961
A;Accession: JQ1366
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             A; Catus: preliminary; nucleic acid sequence not shown; translation not shown
A; Molecule type: DNA
A; Medludes: 1-718 <COL>
A; Colsdudes: 1-718 <COL>
A; Colsdudes: 1-718 <COL>
A; Colsdudes: 1-718 <COL>
A; Experimental source: GB: 292771; GB: AL123456; NID: 93242259; PIDN: CAB07083.1; PID: 91877325
A; Experimental source: strain H37Rv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            C:Superfamily: Enterococcus copper-transporting ArPase copB; ATPase nucleotide-binding F:128-454/Domain: ArPase transduction domain homology <ATT>
F:532-676/Domain: ArPase nucleotide-binding domain homology <ATN>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ö
                                                                                                                                                                            polyprotein - hepatitis C virus (French isolate) (fragments)
C:Species: hepatitis C virus
C:Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 17-Nov-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         C; Accession: T41395
R; Wood, V.; Rajandream, M.A.; Barrell, B.G.; Murphy, L.; Harris, D. Bubnitted to the EMBL Data Library, May 1998
A; Reference number: 221991
A; Accession: T41395
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           75.5%; Score 37; DB 2; Length 718; 63.6%; Pred. No. 63;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Length 716,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 A;Status: preliminary; translated from GB/EMBL/DDBJ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Score 42; DB 2;
Pred. No. 7.6;
2; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Pred. No. 63;
3; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Query Match 85.7%;
Best Local Similarity 72.7%;
Matches 8; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Local Similarity 63.6 les 7; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1 GSVVIVGRIVL 11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1 GSVVIVGRIVL 11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Best Loc
Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  QQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         å
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
```

ô

Gaps

ő

```
GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.
```

OM protein - protein search, using sw model

August 30, 2003, 18:01:52; Search time 0.544792 Seconds (without alignments) 949.524 Million cell updates/sec Run on:

US-09-965-594-26 49 1 GSVVIVGRIVL 11

Title: Perfect score: Sequence:

Scoring table:

BLOSUM62 Gapop 10.0 , Gapext 0.5

127863 seqs, 47026705 residues Searched:

127863 Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0% Maximum Match 100% Listing first 45 summaries

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution. SwissProt\_41:\* Database :

	Description	P26663 h qenome po	h genome		æ	h genome	h denome	mycobacte				equine	equine	eduine			-		0	_	Q9hy84 pseudomonas	_	P11098 escherichia		Q9ccll mycobacteri		••	Q53138 rhodococcus	9	Q8zk55 salmonella	O9ch12 lactococcus	32808	Q,	_
SUMMARIES	ID	POLG_HCVBK	POLG_HCVJA	POLG_HCVJT	POLG_HCVTW	POLG_HCVH	POLG_HCV1	CIPC_MYCTU	DPO2_SCHPO	YMO7_YEAST	MCH_METTR	ENV_EIAV1	ENV_EIAV2	ENV_EIAV3	ENV_EIAV9	ENV_EIAVC	ENV_EIAVW	ENV_EIAVY	ENV_EIAV5	MCH_ARCFU	ASSY_PSEAE	SHU7_ECOLI	CYSG_ECOLI	JI60_HORVU	CTPC_MYCLE		CND1_SCHPO	COBM_RHOER	PLSX_THETN	Y8C8_SALTY	DEOB_LACLA	DEOB_LACIC	HS47_HUMAN	HS47_MOUSE
	Query Match Length DB	3010 1												859 1		859 1	859 1	859 1	860 1	316 1	405 1	456 1	457 1	560 1	725 1	968 1	1158 1	249 1	332 1	388 1	411 1	411 1	417 1	417 1
ď	Query Match	98.0	0.86	98.0		86.8		75.5	71.4	4.69	69.4	4.69	69.4	4.69	69.4	69.4	69.4	4.69	4.69	67.3	67.3	67.3		67.3	67.3	67.3	67.3	2	65.3	'n.	δ.	65.3		65.3
	Score	48	48	4.8	48	44	<b>4</b> 3	37	35	34	34	34	34	34	34	34	34	34	34	33	33	33	33	33	33	33	33	32	32	32	32	32	32	32
	Result No.	7	7	m	*		9	7	80	6	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	56	27	28	29	30	31	. 32	33

P29457 rattus norv		P44023 haemophilus							P95993 sulfolobus		_
HS47_RAT	CBP2_HUMAN	YFCC_HAEIN	CTPB_MYCTU	YB29_YEAST	RL27_CHLTE	YNES_BACSU	VTAI_LAMBD	NEUA_HAEIN	RS2_SULSO	FGFN_RAT	UPKA_BOVIN
1	П	Н	Н	ч	M	-	-	ч	ч	ч	-
417	418	503	752	855	8	193	223	228	231	251	258
65.3	65.3	65.3	65.3	65.3	63.3	63.3	63.3	63.3	63.3	63.3	63.3
32	32	32	32	32	31	31	31	31	31	31	31
34	35	36	37	38	38	40	41	42	43	44	45

## ALIGNMENTS

RESULT 1 POLG_HCYBK ID POLG_HCYBK STANDARD; PRT; 3010 AA.	01-AUG-1992 01-AUG-1992		Envelope glycopro				Hepatitis C virus (isolate BK)	Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae; Hepacivirus			SEQUENCE FROM N.A. MEDLINE-91140698: PubMed-1847440;			from human carriers.";			MEDLINE-96235224; PubMed-	Borowski P., Heiland M., Oehlmann K., Becker B., Korneteky L.,	Non-structural protein 3 of hepatitis C virus inhibits			X-RAY CRYSTALLOGRAPHY (2.4 ANGSTROMS) OF 1027-1215.		MOOMBIN E.W.		trypsin-like fold and a structural zinc binding site."; Cell 87:331-342(1996).		X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS) OF 1027-1210 AND 1678-1691. MEDLINE-98227846: Pubmed-9568891:			virus: a 2.2-A resolution structure in a hexagonal crystal form.			-	precursor	position, Cys or Thr in Pl and Ser or Ala in Pl'!- CATALYTIC ACTIVITY: N nucleoside triphosphate - N diphosphate +
E C C	222	DI F	300	2 2	DE	E E	os	88	88	RN	ž %	\$ 3	χ μ ξ Ε	RT	RL	S C	ΧX	\$	er e	R.L	RN	2 Y	2 2	Æ	RT	RI	KN	7 Y	2	Z Z	E.	2 C	ပ္ပ	ဗ္ဗဗ္ဗ	8	ខម

a

us-09-965-594-26.rsp

```
01-AGC-1992 (Rel. 23, Created)
01-AGC-1992 (Rel. 23, Last sequence update)
28-FEB-2003 (Rel. 41, Last annotation update)
Genome polyprotein (Contains: Capsid protein C (Core protein) (P22);
Envelope 91ycoprotein El (GP32) (GP35); Envelope 91ycoprotein E2
                                                                                                                                                                                                                                                                           (POTENTIAL)
(POTENTIAL)
(POTENTIAL)
(POTENTIAL)
(POTENTIAL)
(POTENTIAL)
                                                                             (POTENTIAL).
(POTENTIAL).
(POTENTIAL).
(POTENTIAL).
(POTENTIAL).
(POTENTIAL).
(POTENTIAL).
(POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Score 48; DB 1; Length 3010;
Pred. No. 1.1;
1; Nismatches 0; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              327189 MW; F8422D5ECCFDFD9C CRC64;
                                                         GICNAC
GGLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      PRT; 3010 AA
                            N-LINKED
N-L
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             98.08;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1678 GSVVIVGRIIL 1688
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Query Match
Best Local Similarity 90.9
Matches 10; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1 GSVVIVGRIVL 11
       L 31

L 31

L 31

L 32

                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              10076
10081
10081
10082
10092
11093
11120
11120
11133
11134
11164
11164
11166
11166
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1186
1188
1197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1189 119
1198 120
1203 120
1680 168
3010 AA;
                                                                                                                                                                                                                                                                                                                                                      POLG_HCVJA
P26662;
                            CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                      CARBOHYD
CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                 CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                    CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
STRAND
                                                                                                                                                                                                       CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                             CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          STRAND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                STRAND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                HELLIX
STRAND
STRAND
STRAND
TURN
STRAND
HELLIX
TURN
STRAND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 STRAND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 STRAND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       STRAND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  RESULT 2
POLG_HCVJA
ID POLG_HC
DT 01-AGG
DT 01-AGG
DT 28-FEB
DT 28-FEB
DE GENOME
DE ENVELO
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        HELIX
TURN
          ô
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                        This SWISS-PROT entry is copyright. It is produced through a collaboration between the SWISS institute of Bioinformatics and the BMBL outstation the European Bloinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                MEROPS; 1939.001; -.

MEROPS; 1949.001; -.

MEROPS; 1949.001; -.

MEROPS; 1949.001; -.

MICHEPTO: IPR001521; HCV_core.

MICHEPTO: IPR001521; HCV_core.

MICHEPTO: IPR001521; HCV_NOST.

MICHEPTO: IPR001521; HCV_NOST.

MICHEPTO: IPR001040; HCV_NOST.

MICHEPTO: IPR00140; HCV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    NONSTRUCTURAL PROTEIN NS4A (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4B (POTENTIAL).
RNA-DIRECTED RNA-POLYMERASE (POTENTIAL).
CHARGE RELAY SYSTEM.
CHARGE RELAY SYSTEM.
CHARGE RELAY SYSTEM.
ATP (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CELLULAR AMINOPEPTIDASE.
CELLULAR AMINOPEPTIDASE.
AMATRIX PROTEIN (FOTENTIAL).
MAJOR ENVELOPE PROTEIN E (POTENTIAL).
MAJOR ENVELOPE PROTEIN NS | POTENTIAL).
NONSTRUCTURAL PROTEIN NS2 (POTENTIAL).
PROTEASE/HELICASE NS3 (POTENTIAL).
                                 K
{RNA}(N).
-!- SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY LIPOPROTEIN ENVELOPE CONSISTS OF TWO PROTEINS: PROTEIN M AND GLYCOPROTEIN B. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MRNA.
-!- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY S29.
                                                                                                                                                                                                                                                                                                                                                                           EMBL; M58335; AAA72945.1; -
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              729
1006
1615
1862
2013
3010
369
1083
1107
1165
                                                                                                                                                                                                                                                                                                                                                                                                  A38465; GNWVTC.

1JANQ: 25-MAR-98.

1JAS: 14-JAN-98.

1NC3: 08-APR-98.

1C2P: 15-NOV-00.

1CSJ: 08-NOV-99.

1GX5: 09-APR-02.

1GX5: 09-APR-02.

1GX6: 10-APR-02.

1GX6: 10-APR-02.

1GX6: 10-APR-02.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              MEROPS; S29.001
MEROPS; U39.001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ACT_SITE
ACT_SITE
ACT_SITE
NP_BIND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    FRANSHEM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CHAIN
CHAIN
CHAIN
CHAIN
CHAIN
CHAIN
CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                       PIR;
PDB;
PDB;
PDB;
PDB;
PDB;
PDB;
PDB;
```

ö

Gaps

٣

```
a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            à
                                                                                                                                                                                                                                                                                                                                                                                        This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its mose by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
(GP68) (GP70) (NS1); Protein P7; Nonstructural protein NS2 (P21) (EC 3.4.22.-); Protease/helicase NS3 (P70) (Hepacivirin) (EC 3.4.21.98); Nonstructural protein NS48 (P4); Nonstructural protein NS48 (P57); Nonstructural protein NS58 (P56); Nonstructural protein NS58 (P56) (P70) (RNA-directed RNA polymerase) (EC 2.7.7.48)]. Hepatitis C virus (Isolate Japanese) (HCV).
                                                                                                                                                                                  [RNA](N).
-1- SUBGUTT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY
LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF
PROTEIN C AND MRNA.
                                                                                                    SEQUENCE FROM N.A.
MEDITNE-91088550; Pubmed-2175903;
Kato N., Hijikata M., Ootsuyama Y., Nakagawa M., Ohkoshi S.,
Sugimura T., Shimotohno K.;
"Modecular cloning of the human hepatitis C virus genome from
"Appanese patients with non-A, non-B hepatitis.":
Proc. Natl. Acad. Sci. U.S.A, 87:9524-9528(1990).
                                                                                                                                                                                                                                                                                                                                                                        -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY S29.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             InterPro; IPR001266; HCV_RGRP.
InterPro; IPR001569; Halicase C.
InterPro; IPR001095; RNA_POL_DS_PS.
InterPro; IPR007094; RNA_pol_PSvir.
Pfam; PP01543; HCV_capsid; l.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         InterPro: IPR001410; DEAD.
InterPro: IPR002522; HCV_capsid.
InterPro: IPR002521; HCV_corc.
InterPro: IPR002519; HCV_corc.
InterPro: IPR002519; HCV_NS1.
InterPro: IPR002518; HCV_NS2.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Interpro; IPR004109; HCV_NS3.
Interpro; IPR001409; HCV_NS4a.
Interpro; IPR001490; HCV_NS4b.
Interpro; IPR002868; HCV_NS5a.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                         EMBL; D90208; BAA14233.1; -. PIR; A39253; GNWVCJ. HSSP; P26663; LJXP. MEROPS; S29.001; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           HCV_NS3; 1.
HCV_NS4a; 1.
HCV_NS4b; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  HCV_core; 1
HCV_env; 1.
HCV_NS1; 1.
HCV_NS2; 1.
                                                                                  NCBI_TaxID-11116;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               PF01001;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Pfam;
Pfam;
```

```
Pfam; PF01506; HCV_NS5a; 1.

R Pfam; PF00291; helicase_C; 1.

R Promom; PD186062; HCV_NS1; 1.

R Probom; PD186062; HCV_NS1; 1.

R SWART; SM0487; DEXDC; 1.

W Core protein; Glycoprotein; Transferase; RNA-directed RNA polymerase; Core protein; Glycoprotein; Bruvelope protein; Helicase; ATP-binding; Transmembrane; Nonstructural protein; Hydrolase; Serine protease.

FT CARIN 116 191 HALOR PROTEIN (POTENTIAL).

FT CHAIN 116 191 HALOR ENVELOR PROTEIN (EOTENTIAL).

FT CHAIN 118 192 383 HALOR ENVELOR PROTEIN NS1 (POTENTIAL).

FT CHAIN 1006 NON-STRUCTURAL PROTEIN NS2 (POTENTIAL).

FT CHAIN 1016 NON-STRUCTURAL PROTEIN NS4 (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             01-APR-1993 (Rel. 25, Created)
01-APR-1993 (Rel. 25, Last sequence update)
15-SEP-2003 (Rel. 42, Last annotation update)
15-SEP-2003 (Rel. 42, Last annotation update)
66nome polyprotein (Contains: Capsid protein C (Core protein) (P22);
Envelope glycoprotein E1 (GP32) (GP35); Envelope glycoprotein E2 (GP68) (GF70) (NS1); Protein P7: Nonstructural protein NS2 (P21)
(EC 3.4.22.-); Protease/helicase NS3 (P70) (Hepacivirin)
(EC 3.4.21.98); Nonstructural protein NS4A (P4); Nonstructural protein NS5B (P27); Nonstructural protein NS5B (P66) (P70) (RNA-directed RNA Polymerase) (EC 2.7.7.48)].
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae; Hepacivirus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   RELAY SYSTEM (BY SIMILARITY).
RELAY SYSTEM (BY SIMILARITY).
RELAY SYSTEM (BY SIMILARITY).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       0;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        (POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Score 48; DB 1; Length 3010;
Pred. No. 1.1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            0; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   MW; AA993794F46DB185 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (GLCNAC...) (GLCNAC...) (GLCNAC...) (GLCNAC...) (GLCNAC...) (GLCNAC...) (GLCNAC...) (GLCNAC...) (
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      N-LINKED (GLCNAC...)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  (GLCNAC. . .)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       (GLCNAC. . .)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      N-LINKED (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 N-LINKED (GLCNAC
N-LINKED (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ATP (POTENTIAL).
DECH BOX.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                             POTENTIAL.
CHARGE REL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Pred. No
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CHARGE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CHARGE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                327017
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          98.0%;
90.9%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Query Match
Best Local Similarity 90.97
....nhes 10; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1678 GSVVIVGRIIL 1688
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1 GSVVIVGRIVL 11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 2240
2788
                                                                                                                                                                                                                                                                                                                    729
1006
1615
1862
2013
2013
369
1083
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Ž
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    NCBI_TaxID-31642;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        POLG_HCVJT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CARBOHYD
CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CARBOHYD
CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ACT_SITE
NP_BIND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                               FRANSMEN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ACT_SITE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      RESULT 3
POLG_HCVJT
```

```
NCBI_TaxID=31645;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Hepacivirus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             POLG_HCVTW
                                                                        ACT_SITE
ACT_SITE
ACT_SITE
NP_BIND
                                                                                                                                                                                                                                                                                                                            CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                       CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                CARBOHYD
                                                                                                                                                                                                                                                                        CARBOHYD
                                                                                                                                                                                                                                                                                      CARBOHYD
                                                                                                                                                                                                                                                                                                     CARBOHYD
                                                                                                                                                                                                                   CARBOHYD
                                                                                                                                                                                                                                           CARBOHYD
                                                                                                                                                                                                                                                           CARBOHYD
                                                                                                                                                                                                                                                                                                                 CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                     CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                 CARBOHYD
                                                                                                                                              CARBOHYD
                                                                                                                                                            CARBOHYD
                                                                                                                                                                           CARBOHYD
                                                                                                                                                                                       CARBOHYD
                                                                                                                                                                                                    CARBOHYD
                                                                                                                                                                                                                                CARBOHYD
                                                           TRANSMEM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            P29846;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  RESULT 4
POLG_HCVTW
                                              CHAIN
    a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            õ
                                                                                                                                                                                                                                                                                                         This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss institute of Bioinformatics and the EMEL outstation the European Bloinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
       Tanaka T., Kato N., Nakagawa M., Ootsuyama Y., Cho M.J.,
Tanaka T., Kato N., Nakagawa M., Ootsuyama Y., Cho M.J.,
Tanaka T., Kijikata M., Ishimura Y., Shimotohno K.;
"Molecular cloning of hepatitis C virus genome from a single Japanese
"arriter: sequence variation within the same individual and among
"infected individuals.";
"Infected individuals.";
"Infected individuals.";
"Sarriter: Sequence variation within the same individual and among
"I. VIUS RES. 23:39:34(192).
"INSTANCE TO SECTION OF PROSEIBLE MEMBRANE-RELATED FUNCTION.
NSS AND NSS MAY PLAY A ROLE IN THE VIRAL RNA REPLICATION.
"C -1 CATALYTIC ACTIVITY: Hydrolysis of four peptide bonds in the viral
precursor polyprotein, commonly with Asp or Glu in the P6
position, Cys or Thr in P1 and Ser or Ala in P1'.
"C -1 CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CELLULAR AMINOPEPTIDASE.
CAPSID PROTEIN C (POTENTIAL).
MATRIX PROTEIN (POTENTIAL).
MAJORULOPE PROTEIN E (POTENTIAL).
NONSTRUCTURAL PROTEIN NSI,E2 (POTENTIAL).
NON STRUCTURAL PROTEIN NSI,E2 (POTENTIAL).
                                                                                                                                                                                                                         -i-SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS: PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MRNA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Polyprotein; Glycoprotein; Transferase; RNA-directed RNA polymerase; Core protein; Coat protein; Envelope protein; Helicase; ATP-binding; Transmembrane; Nonstructural protein; Hydrolase; Serine protease;
                                                                                                                                                                                                                                                                                -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY $29
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    REMOVED FROM CAPSID PROTEIN C
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     InterPro; 19701410; DEAD.
InterPro; IPR001221; HVV_capsid.
InterPro; IPR002512; HVV_capsid.
InterPro; IPR002513; HVV_capsid.
InterPro; IPR002519; HVV_core.
InterPro; IPR002519; HVV_NS1.
InterPro; IPR001409; HVV_NS3.
InterPro; IPR001409; HVV_NS3.
InterPro; IPR001409; HVV_NS4.
InterPro; IPR001409; HVV_NS4.
InterPro; IPR001409; HVV_NS4.
InterPro; IPR001045; HVV_RS4.
InterPro; IPR001045; HVV_RS4.
InterPro; IPR001094; RNA_DOL_DS_PS.
InterPro; IPR001094; RNA_DOL_DS_PS.
InterPro; IPR001094; RNA_DOL_PSvir.
Pfam; PP01540; HCV_capsid; I.
Pfam; PP01540; HCV_capsid; I.
Pfam; PP01560; HCV_NS1; I.
Pfam; PP01006; HCV_NS1; I.
Pfam; PP01010; HCV_NS4; I.
Pfam; PP01006; HCV_NS4; I.
Pfam; PP01006; HCV_NS4; I.
Pfam; PP01006; HCV_NS54; I.
Pfam; PP01006; HCV_NS54; I.
Pfam; PP01006; HCV_NS54; I.
Pfam; PP01006; HCV_NS54; I.
Pfam; PP01007; HCV_NS54; I.
                                                                                                                                                                                                                                                                                                                                                                                                                         EMBL; D11168; BAA01943.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               115
191
383
729
1006
                                                                                                                                                                                                                                                                                                                                                                                                                                     PIR; A45573; A45573.
PDB; 1A1Q; 25-MAR-98.
PDB; 1JXP; 14-JAN-98.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                             MEROPS; 529.001
MEROPS; U39.001
                                                                                                                                                                                                              {RNA}(N).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   3D-structure.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CHAIN
CHAIN
CHAIN
```

```
.:
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             *** PEDLINE-92230206; PubMed-1314449;

*** MEDLINE-92230206; PubMed-1314449;

*** MEDLINE-92230206; PubMed-1314449;

*** Chen P.J., Lin M. H., Tal K.F., Liu P.C., Lin C.J., Chen D.S.;

*** Chen P.J., Lin M. H., Tal K.F., Liu P.C., Lin C.J., Chen D.S.;

*** Chen P.J., Lin M. H., Tal K.F., Liu P.C., Lin C.J., Chen D.S.;

*** The Taiwanese hepatitis C viral genomic and antigenomic RNA.*;

*** The Taiwanese hepatitis C viral genomic and antigenomic RNA.*;

*** The Taiwanese hepatitis C viral genomic and antigenomic RNA.*;

*** The Taiwanese hepatitis C viral genomic and antigenomic RNA.*;

*** The Taiwanese hepatitis C viral genomic and antigenomic RNA.*;

*** The Taiwanese hepatitis C viral genomic RNA.*;

*** The Taiwanese hepatitis C vir
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     01-APR-1993 (Rel. 25, Created)
01-APR-1993 (Rel. 25, Last sequence update)
15-SEP-2003 (Rel. 25, Last sequence update)
66-00me polyprotein [Contains Capsid protein C (Core protein) (P22);
67-2003 (Rel. 20, Last annotation update)
67-2003 (Rel. 20, Last annotation update)
67-2004 (MSI); Protein P7: Nonstructural protein RS (GPR) (GPV) (MSI); Protease/helicase NSJ (P70) (Hepacivirin)
62-3-4.21.89; Nonstructural protein NSAA (P4); Nonstructural protein
NSAB (P27): Nonstructural protein NSAA (P4); Nonstructural protein
NSAB (P60) (P70) (RNA-directed RNA polymerase) (EC 2.77.48)].
Hepatitis C virus (isolate Taiwan) (HCV).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       the viral
PROTEASE/HELICASE NS3 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS44 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS48 (POTENTIAL).
RNA-DIRECTED RNA POLYMERASE (POTENTIAL).
                                                                                                                                                                                      SIMILARITY).
SIMILARITY).
SIMILARITY).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                (POTENTIAL)
(POTENTIAL)
(POTENTIAL)
(POTENTIAL)
(POTENTIAL)
(POTENTIAL)
(POTENTIAL)
(POTENTIAL)
(POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ó
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Score 48; DB 1; Length 3010;
Pred. No. 1.1;
1; Mismatches 0; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                NKED (GLCNAC. . .) (POTI 94A1C77435D642BB CRC64;
                                                                                                                                                                                      (BY
(BY
(BY
                                                                                                                                                                                  CHARGE RELAY SYSTEM
CHARGE RELAY SYSTEM
CHARGE RELAY SYSTEM
                                                                                                                                                                                                                                                                                                                                                                                                         N-LINKED GECNAC
                                                                                                                                                                                                                                                                                                      ATP (POTENTIAL).
DECH BOX.
                                                                                                                                                                                                                                                                                                                                                                                 N-LINKED (GLCNAC
                                                                                                                                                          POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    326573
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            98.08;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Query Match
Best Local Similarity 90.9"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1678 GSVVIVGRIIL 1688
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1 GSVVIVGRIVL 11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     2529 253
2788 278
3010 AA;
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMEL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb.sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     InterPro: IPR002868; HCV_RASE.

N InterPro: IPR002868; HCV_RASE.

N InterPro: IPR00106; HCV_RASE.

N InterPro: IPR001095; RNA_POI_DS_PS.

InterPro: IPR001094; RNA_POI_DS_PS.

InterPro: IPR001094; RNA_POI_DS_PS.

InterPro: IPR001094; RNA_POI_DS_PS.

N Pfam; PP01542; HCV_cors; 1.

N Pfam; PF01539; HCV_ROS; 1.

N Pfam; PF01539; HCV_ROS; 1.

N Pfam; PF01000; HCV_NS; 1.

N Pfam; PF01000; HCV_NS; 1.

N Pfam; PF01000; HCV_NS; 1.

N Pfam; PF00271; HCV_NS; 1.

N Pfam; PF00271; HCV_NS; 1.

N Pfam; PF00398; Viral_RARE; 1.

N Pfam; PF00998; Viral_RARE; 1.

N Probom; PD0998; Viral_RARE; 1.

N Probom; PD186062; HCV_NS; 1.

N Probom; PD186062; H
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CELLULAR AMINOPEPTIDASE.
CORE PROTEIN (POTENTIAL).
MAJOR PROTEIN (POTENTIAL).
MAJOR ENVELOPE PROTEIN E (POTENTIAL).
NONSTROCTURAL PROTEIN NS1 (POTENTIAL).
PROTEASE/HELICASE NS3 (POTENTIAL).
NONSTROCTURAL PROTEIN NS1 (POTENTIAL).
NONSTROCTURAL PROTEIN NS4 (POTENTIAL).
NONSTROCTURAL PROTEIN NS48 (POTENTIAL).
RNA-DIRECTED RNA POLYMERASE (POTENTIAL).
                              SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY LIPPOFROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS: PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C. AND MRNA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           POTENTIAL.
CHARGE RELAY SYSTEM (BY SIMILARITY).
CHARGE RELAY SYSTEM (BY SIMILARITY).
CHARGE RELAY SYSTEM (BY SIMILARITY).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  REMOVED FROM CAPSID PROTEIN C BY THE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (POTENTIAL).
(POTENTIAL).
(POTENTIAL).
(POTENTIAL).
(POTENTIAL).
                                                                                                                                                SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY S29.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           (GLCNAC. . .)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ATP (POTENTIAL).
DECH BOX.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 N-LINKED (N-LINKED (N-LINK
                                                                                                                                                                                                                                                                                                                                                                                                                        EMBL; M84754; -; NOT_ANNOTATED_CDS
PIR; A40244; GNWVTW.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Interpro; IPR002519; HCV_env.
Interpro; IPR002511; HCV_env.
Interpro; IPR002511; HCV_NS1.
Interpro; IPR004109; HCV_NS2.
Interpro; IPR0004109; HCV_NS3.
Interpro; IPR00149; HCV_NS4.
Interpro; IPR001499; HCV_NS4a.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  IPR002522; HCV_capsid.
IPR002521; HCV_core.
IPR002519; HCV_env.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                InterPro; IPR001410; DEAD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            729
1006
1006
1006
3010
3010
3010
1003
1105
1105
1106
233
233
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1N64; 25-FEB-03,
1N53; 08-APR-98,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    MEROPS; S29 001;
MEROPS; U39.001;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   3D-structure.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ACT_SITE
ACT_SITE
ACT_SITE
NP_BIND
SITE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CARBOHYD
CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           TRANSMEM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CHAIN
CHAIN
CHAIN
CHAIN
CHAIN
CHAIN
```

```
01-40G-1992 (Rel. 23, Created)
01-40G-1992 (Rel. 23, Last sequence update)
15-5EP-2003 (Rel. 24, Last annotation update)
66-00me polyprotein (Contains Capsid protein C (Core protein) (P22);
670-63) (GP70) (NS1); Protein P1 (GP32) (GP35); Envelope glycoprotein E2 (GP48) (GP32) (GP35); Envelope glycoprotein E2 (GP48) (GP30) (NS1); Protein P2 Nonstructural protein NS4 (P499-1); Protease/helicase A33 (P70) (Hepacivirin) (EC 3.4.21.98); Nonstructural protein NS5A (P65); Nonstructura
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        X-RAY CRYSTALLOGRAPHY (2.1 ANGSTROMS) OF 1207-1657.
MEDLINE-9731312; PubMed-9187654;
Ydo N., Hesson T., Cable M., Hong Z., Kwong A.D., Le H.V., Weber P.C.;
Scructure of the hepatitis C virus RNA helicase domain.";
Nat. Struct. Biol. 4:463-467(1997).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Kim J.L., Morgenstern K.A., Griffith J.P., Dwyer M.D., Thomson J.A., Murcko M.A., Lin C., Caron P.R.;
"Hepatitis C virus NS3 RNA helicase domain with a bound oligonucleotide: the crystal structure provides insights into the mode of unwinding.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  OLD THE CLEAVAGE OF NS2-NS3.

STRUCTION: PROTEASE NS2 IS RESPONSIBLE FOR THE CLEAVAGE OF NS2-NS3.

-I- FUNCTION: PROTEASE NS3 IS RESPONSIBLE FOR THE CLEAVAGE OF NS3-NS3-NS3-NS4A, NS4A-NS4B, NS4B-NS5A AND NS5A-NS5B.

-I- FUNCTION: NS4A-FORMS A COMPLEX WITH NS3 AND IS ESSENTIAL FOR THE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   "Genomic structure of the human prototype strain H of hepatitis C virus: comparison with American and Japanese isolates."; Proc. Natl. Acad. Sci. U.S.A. 88:10292-10296(1991).
                                                                                              (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                    (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                            (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ;
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 MEDLINE-92052256; PubMed-1658800;
Inchauspe G., Zebedee S., Lee D.H.H., Sugitani M., Nasoff M.,
| 305 | N-LINKED (GLCNAC...) (POTENT ALINKED (GLCNAC...) (POTENT ALINKED (GLCNAC...) (POTENT ALINKED (GLCNAC...) (POTENT N-LINKED (GLCNAC...)) (POTENT N-LINKED (GLCNAC...) (POTENT N-LINKED (GLCNAC...) (POTENT N-LINKED (GLCNAC...)) (POTENT N-LINKED (GLCN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Score 48; DB 1; Length 3010;
Pred. No. 1.1;
1; Mismatches 0; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS) OF 1192-1657.
WEDLINE-98154321; Pubmed-9493270;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                PRT; 3011 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           98.0%;
90.9%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Ouery Match
Best Local Similarity 90.9
Matches 10; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1 GSVVIVGRIVE 11
      305
417
423
430
448
532
532
540
556
623
645
645
                                                                                                                                                                                                                                                                                                                                                                                       2240
2529
2788
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               NCBI_TaxID=11108;
      Prince A.H.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Hepacivirus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                POLG_HCVH
                                  CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
                                                                                                                                                       CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                     CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                 CARBOHYD
                                                                                                                                                                                                                   CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                            SEQUENCE
                                                                                                                                                                                                                                               CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                          CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               P27958
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       RESULT S
POLG_HCVH
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ô
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  qq
```

9

```
SIMILARITY).
SIMILARITY).
SIMILARITY).
                                                                                                             (POTENTIAL).
    NONSTRUCTURAL PROTEIN NSSA.
                                            (BY
(BY
                                         CHARGE RELAY SYSTEM CHARGE RELAY SYSTEM CHARGE RELAY SYSTEM
                                                                                                           N-LINKED (GLCNAC, N-LINKED GLCNAC, N-LIN
                                                                                   ATP (POTENTIAL).
DECH BOX.
                                                                                                                                                                                                                                                                                         N-LINKED
N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1319
1335
1340
1347
1353
1361
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1380
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1385
1393
1409
1411
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1417
1420
1436
1439
1453
1463
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1480
1488
1490
1501
1507
                                                                                                                                                                                1343
1352
1362
1363
1373
1376
1376
1376
1389
1389
1389
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1414
1419
1432
1438
1450
                                                                                                                                                                                                                                                                                                                                                                                        2551
272
272
272
272
273
273
302
315
315
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            317
323
336
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
                                                                                                                                                                                                          CARBOHYD
CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                 CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
STRAND
TURN
HELLIX
TURN
                                           ACT_SITE
ACT_SITE
ACT_SITE
ACT_SITE
NP_BIND
SITE
                                                                                                                             CARBOHYD
                               TRANSMEM
                                                                                                              CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                               STRAND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    HELIX
TURN
STRAND
TURN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 TURN
STRAND
HELIX
STRAND
HELIX
TURN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                STRAND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            STRAND
HELIX
TURN
STRAND
STRAND
STRAND
HELIX
HELIX
                                                                                                                                                                                                                                                                                                                                                                                                                                                          STRAND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            STRAND
STRAND
STRAND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           STRAND
TURN
STRAND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  STRAND
                                                                                                                                                                                                                                                                                                                                                                                        STRAND
HELIX
TURN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     STRAND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            STRAND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        HELIX
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               TURN
    This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the Buropaen Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
            FUNCTION: NS5B IS A RNA-DEPENDENT RNA POLYMERASE THAT FLAYS AN ESSENTIAL ROLE IN THE VIRUS REPLICATION.

CATALYTIC ACTIVITY: Hydrolysis of four peptide bonds in the viral precursor polyprotein, commonly with Asp or Glu in the P6 position, Cys or Thr in P1 and Ser or Ala in P1'.

CATALYTIC ACTIVITY: N nucleoside triphosphate +
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Polyprotein; Glycoprotein; Transferase; RNA-directed RNA polymerase;
Core protein; Coat protein; Envelope protein; Helicase; ATP-binding;
Transmembrane; Nonstructural protein; Hydrolase; Serine protease;
                                                                                                        SUBDINIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS: E1 AND E2. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MRNA.

PTH: THE STRUCTURAL PROTEINS C. E1 AND E2 ARE PRODUCED BY PROTEOLYTIC PROCESSING BY THE HOST SIGNAL PEPTIDASES.

SIMILARITY: THE NS2 PROTEASE BELONGS TO PEPTIDASE FAMILY 039.

SIMILARITY: THE NS3 PROTEASE BELONGS TO PEPTIDASE FAMILY S29.
  NS5A SEEMS TO HAVE A TRANSCRIPTIONAL ACTIVATORY ROLE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      REMOVED FROM CAPSID PROTEIN C BY THE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  PROTEASE/HELICASE NS3.
NONSTRUCTURAL PROTEIN NS4A.
NONSTRUCTURAL PROTEIN NS4B.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      NONSTRUCTURAL PROTEIN NS2.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ENVELOPE GLYCOPROTEIN E2.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CELLULAR AMINOPEPTIDASE.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CAPSID PROTEIN C
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         PROTEIN P7
                                                                                                                                                                                                                                                                                                                                                                                                                         TRANSFAC; TO4155; ...
InterPro: IPR001410; DEAD.
InterPro: IPR001521; HCV_capsid.
InterPro: IPR001521; HCV_capsid.
InterPro: IPR001519; HCV_capsid.
InterPro: IPR001519; HCV_ms1.
InterPro: IPR001519; HCV_ms2.
InterPro: IPR001490; HCV_ms3.
InterPro: IPR001490; HCV_ms3.
InterPro: IPR001490; HCV_ms4.
InterPro: IPR001650; HCV_ms5a.
InterPro: IPR001650; HCV_ms1; I.
Pfam; PF011543; HCV_ms1; I.
Pfam; PF011560; HCV_ms1; I.
Pfam; PF01160; HCV_ms1; I.
Pfam; PF01006; HCV_ms3; I.
Pfam; PF01006; HCV_ms3; I.
Pfam; PF01006; HCV_ms3; I.
Pfam; PF01006; HCV_ms3; I.
Pfam; PF01071; Helicase_C; I.
Pfam; PF00271; Helicase_C; I.
Probm; PP1086062; HCV_ms1; I.
                                                                                                                                                                                                                                                                                                                          EMBL; M67463; AAA45534.1; -. PIR; A36814; GNWVCH. PDB: 1HEI; 25-NOV-98. PDB: 1AlV; 16-FEB-99. PDB: 1AlV; 17-UN-98. MEROPS; U39.001; -. MEROPS; U39.001; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      3D-structure.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CHAIN
CHAIN
CHAIN
CHAIN
CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CHAIN
```

```
Pfam;
Pfam;
Pfam;
Pfam;
                                                                                                                                                                                                                                                                                                                                                                       Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Best Loc
Matches
                 This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bloinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entitles requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                 ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               MEDLINE—91172826; PubMed=1848704;

MEDLINE—91172826; PubMed=1848704;

MEDLINE—91172826; PubMed=1848704;

Choo Q.-L., Richman K.H., Han J.H., Berger K., Lee C., Dong C.,

A Gallegos C., Coit D., Medina-Sebby A., Barr P.J., Weiner A.J.,

Bradley D.W., Kuo G., Houghton M.;

"Genetic organization and diversity of the hepatitis C virus.";

"Genetic organization and diversity of the hepatitis C virus.";

"Genetic organization and diversity of the hepatitis C virus.";

"Genetic organization and diversity of the hepatitis C virus.";

"Genetic organization and diversity of the hepatitis C virus.";

"Genetic organization and diversity of four peptide bonds in the viral precursor polyprotein, commonly with Asp or all in the P6 position, Cys or Thi in pland Ser or Ala in pl. Containing Contain
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        01-806-1992 (Rel. 23, Created)
01-806-1992 (Rel. 23, Last sequence update)
01-806-1992 (Rel. 23, Last sequence update)
01-806-1992 (Rel. 42, Last annotation update)
Genome polyprotein (Contains: Capsaid protein C (Core protein) (P22);
Envelope glycoprotein El (GP32) (GP35); Envelope glycoprotein E2 (GP68) (GP36) (GR15); Envelope glycoprotein E2 (GP68) (GR16) (MS1); Protease/hellcase MS3 (P70) (Hepacivirin) (EC 3.4.22.-); Protease/hellcase MS3 (P70) (Hepacivirin) (EC 3.4.21.98); Nonstructural protein NSSA (P4); Nonstructural protein NSSA (P65); Nonstructural protein NSSA (P66) (P70) (RNA-directed RNA polymerase) (EC 2.7.7.48)].
Hepatitis C virus (isolate I) (HCV)
                                                                                                                                                                                                                                                                                                                 Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              (RNA)(N).
SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS: PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MRNA.
                                                                                                                                                                                                                                                                                                                 ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY S29.
                                                                                                                                                                                                                                                       Score 44; DB 1; Length 3011; Pred. No. 5.7;
1598 1598
1606 1611
1622 1623
1627 1627
1635 1637
1640 1652
3011 AA; 327142 MW; 772CBB29CCD94753 CRC64;
                                                                                                                                                                                                                                                                                                                 1; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               PRT; 3011 A.A.
                                                                                                                                                                                                                                                                                                              0; Mismatches
                                                                                                                                                                                                                                                       88.88;
90.98;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 EMBL; M62321; AAA45676.1; -
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     InterPro; IPR001410; DEAD.
                                                                                                                                                                                                                                                                                                                                                                                              1678 GCVVIVGRIVL 1688
                                                                                                                                                                                                                                                                                                                 Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               STANDARD;
                                                                                                                                                                                                                                                                                                                                                                  1 GSVVIVGRIVL 11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        PDB; 1AlV; 16-FEB-99.
PDB; 1HEI; 25-NOV-98.
                                                                                                                                                                                                                          Ouery Match
Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           MEROPS; S29.001; ...
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           GNWVC3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    NCBI_TaxID-11104;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Hepacivirus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          POLG_HCV1
P26664;
                                                         TURN
STRAND
STRAND
STRAND
HELIX
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   POLG_HCV1
     EELLLES
SHIHHH
                                                                                                                                                                                                                                                                                                                                                                                                                     윱
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               HID DESCRIPTION OF COLUMN OF THE PROPERTY OF THE PROPERTY OF COLUMN OF THE PROPERTY OF THE PRO
```

```
ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     REMOVED FROM CAPSID PROTEIN C BY THE CELLULAR AMINOPEPIDDSE.
CAPSID PROTEIN C (POTEWIIAL).
MATRIX PROTEIN (POTEWIIAL).
MAJOR ENVELOPE PROTEIN E (POTEWIIAL).
NONSTRUCTURAL PROTEIN NSI, POTEWIIAL).
PROTEASE/HELICASE NS3 (POTEWIIAL).
NONSTRUCTURAL PROTEIN NS2 (POTEWIIAL).
NONSTRUCTURAL PROTEIN NS4 (POTEWIIAL).
NONSTRUCTURAL PROTEIN NS4 (POTEWIIAL).
RNA-DIRECTED RNA POLYMERASE (POTEWIIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Polyprotein; Glycoprotein; Transferase; RNA-directed RNA polymerase; Core protein; Coat protein; Envelope protein; Helicase; ATP-binding; Transmembrane; Nonstructural protein; Hydrolase; Serine protese;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CHARGE RELAY SYSTEM (BY SIMILARITY)
CHARGE RELAY SYSTEM (BY SIMILARITY)
CHARGE RELAY SYSTEM (BY SIMILARITY)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (POTENTIAL).
(POTENTIAL).
(POTENTIAL).
(POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Score 43; DB 1; Length 3011;
Pred. No. 8.7;
1; Mismatches 1; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                65F8C9447FCE5AF9 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       (GLCNAC. . .
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             N-LINKED (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (GLCNAC.
(GLCNAC.
(GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (GLCNAC.
(GLCNAC.
(GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ATP (POTENTIAL).
DECH BOX.
LAKON 166; HCV NS5a.

LAKON 166; HCV RGAB.

LATERPO: IPRON 169; RNA_POL_DS_PS.

THERPTO: IPRON 7094; RNA_POL_DS_PS.

THE PED 154; HCV_CORPSIN.

THE PED 159; HCV_CORPSIN.

PRON 159; HCV_CORPSIN.

PRO
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    N-LINKED
                                                                                                                                                                                                                                                                                     Pfam; PF01560; HCV_NS1; 1.
Pfam; PF0158; HCV_NS2; 1.
Pfam; PF01905; HCV_NS2; 1.
Pfam; PF01006; HCV_NS4s; 1.
Pfam; PF01001; HCV_NS4s; 1.
Pfam; PF01506; HCV_NS4b; 1.
Pfam; PF00271; helicase_C; 1.
Pfam; PF00998; Viral_RGRP; 1.
PF0D0m; PD186052; HCV_NS1; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         87.8%;
81.8%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             327197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          115
191
383
729
1006
1615
1862
2013
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SMART; SM00487; DEXDC;
                                                        InterPro; IPR002868;
InterPro; IPR002166;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         3D-structure.
INIT_MET
                                                                                                                   InterPro;
                                                     interPro;
                                                                                                                                            InterPro;
                            InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ACT_SITE
ACT_SITE
ACT_SITE
ACT_SITE
NP_BIND
SITE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             PRANSMEM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CARBOHYD
```

ထ

ó

```
Transmembrane; Phosphorylation; Magnesium; ATP-binding
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           DP02_SCHP0
074946;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SPCC553.09C
  Hydrolase;
                                                                        TRANSMEM
TRANSMEM
                                                                                                                                                                                                                            SEQUENCE
                                                                                                                         TRANSMEM
                                                 TRANSMEM
                                                                                                                                                   MOD_RES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         DPO2_SCHPO
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        XXIIIIIXXX
                                                                                                                                                                                                                                                                                                                                                                           ó
                                                                                                                                                                                                                                                                                                                                                                                                                       a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 between the Swiss Institute of Bioinformatics and the EMBL outstation—the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@lsb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               This SWISS-PROT entry is copyright. It is produced through a collaboration
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          "Whole genome comparison of Mycobacterium tuberculosis clinical and laboratory strains."; Submitted (APR-2001) to the EMBL/GenBank/DDBJ databases.
-!- CATALYTIC ACTIVITY: ATP + H(2)0 - ADP + phosphate.
-!- SUBCELLULAR LOCATION: Integral membrane protein.
-!- SIMILARITY: Belongs to the cation transport ATPases family (P-type ATPases). Subfamily IB.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              STRAIN=CDC 1551 / Oshkosh;
Fleischmann R.D., Alland D., Elsen J.A., Carpenter L., White O.,
Peterson J., DeBoy R., Dodson R., Gwinn M.L., Haft D., Hickey E.,
Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D., Salzberg S.L.,
Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        MEDLINE-9829587; PubMed-9634230; Garnier T., Churcher C., Harris D., Gordon S.V., Eiglmeler K., Gas S., Barry C.E. III, Tekaia F., Gordon S.V., Eiglmeler K., Gas S., Barry C.E. III, Tekaia F., Bordon S.V., Eiglmeler K., Eiglmeler K., Chillingworth T., Connor R., Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holroyd S., Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L., Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L., Rutter S., Geoger K., Selandream M.A., Rajandream M.A., Rogers J., Rutter S., Geoger K., Skelton S., Squares S., Squares R., Shlston J.E., Taylor K., Whitehead S., Barrell B.G.; Peclihering the biology of Mycobacterium tuberculosis from the Nature 393:537-544(1998).
                                                                                                                                                                                                                                                                                                                                            Mycobacterium tuberculosis.
Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
                                                                                                                                                                                                               30-MAY-2000 (Rel. 39, Created)
30-MAY-2000 (Rel. 39, Last sequence update)
31-SEP-20003 (Rel. 42, Last tanoctation update)
Probable cation-transporting P-type APPase C (EC 3.6.3.-).
CTPC OR RV3270 OR MT3370 OR MTCY71.10.
                                                                                                                                                                                                                                                                                                                                                                                          Corynebacterineae; Mycobacteriaceae; Mycobacterium
NCBI_TaxID=1773;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Interpro: IPR006404; Heavy_mct_ATPase.
Interpro: IPR005834; Hydrolase.
Pfam; PF00122; El-E2_ATPase; 1.
Pfam; PF00702; Hydrolase; 1.
PRINTS; PR00119; CATATPASE.
TIGRRAMs; TIGR0151; ATPase-IB2_Cd; 1.
TIGRRAMs; TIGR01525; ATPase-IB_hvy; 1.
TIGRRAMs; TIGR01494; ATPase_P-type; 3.
PROSITE; PS00134; ATPASE_El_E2; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   InterPro; IPR006416; ATPase-IB_hvy.
InterPro; IPR001757; ATPase_E1-E2.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               EMBL; Z92771; CAB07083.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   EMBL, AE007146; AAK47711.1;
PIR; G70978; G70978.
TIGR; MT3370; -.
TubercuList; Rv3270; -.
                                                 1678 GCVVIVGRVVL 1688
                                                                                                                                                                           STANDARD;
1 GSVVIVGRIVL 11
                         SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SEQUENCE FROM N.A.
                                                                                                                                                                        CTPC_MYCTU
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Bishai W.
                                                                                                                                                                                               P96875;
                                                                                                                         RESULT 7
                                                                                                                                                                           DER RESERVE DER RESERVE DE PROPERTIE DE PROP
```

```
*** STEANNSY: A ST
                                                                                                                                                                                                                                                                                                                                                                                                                                                               Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SIMILARITY).

SUBUNIT: DNA POLYMERASE ALPHA-PRIMASE IS A FOUR SUBUNIT ENZYME (SUBUNITS A, B., C AND D), WHICH IS ASSEMBLED THROUGHOUT THE CELL CYCLE. THE LARGEST SUBUNIT (SUBUNITA) HAS DNA POLYMERASE ACTIVITY, THE TWO SMALLEST SUBUNITS (SUBUNITS C AND D) HAVE DNA PRIMASE ACTIVITY. SUBUNIT B BINDS TO SUBUNIT A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              NATURE 415:871-880(2002).
-:- FUNCTION: MAY PLAY AN ESSENTIAL ROLE AT THE EARLY STAGE OF CHROMOSOMAL DNA REPLICATION BY COUPLING THE POLYMERASE ALPHA/PRIMASE COMPLEX TO THE CELLULAR REPLICATION MACHINERY (BY
                                                                                                                                                                                                                                                                                                                                                                                                                                                               ;
0
                                                                                                                                                                                       SIMILARITY).
                                                                                                                                                                                                                                                                                                                                                                                  Score 37; DB 1; Length 718; Pred. No. 30;
                                                                                                                                                                                                                                                                                                                                                                                                                                                               1; Indels
                                                                                                                                                                              PHOSPHORILATION (BY SIMILAR MAGNESIUM (BY SIMILARITY).
MAGNESIUM (BY SIMILARITY).
85D6C93AFE636315 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Schizosaccharomyces pombe (Fission yeast).
Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
Schizosaccharomycetales; Schizosaccharomycetaceae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  15-JUL-1999 (Rel. 38, Created)
15-JUL-1999 (Rel. 38, Last sequence update)
15-EPB-2003 (Rel. 41, Last annotation update)
Probable DNA polymerase alpha subunit B.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 574 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                               3; Mismatches
                                                                         POTENTIAL. POTENTIAL.
                                                                                                                                                  POTENTIAL
                                                                                                                                                                                                                                                                                                        76495 MW;
                                                                                                                                                                                                                                                                                                                                                                                      75.58;
                                                                                                                                                                                                                                                                                                                                                                                                                          63.68;
                                                                                                                                                                                                                                                                                                                                             Ouery Match
Best Local Similarity 63...
Triconservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     286 GSVVVRGRVVV 296
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1 GSVVIVGRIVL 11
                                                                                                                                                                          408
610
614
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Schizosaccharomyces.
NCBI_TaxID=4896;
                                                                                                                                                                                                                                                                614 (718 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SEQUENCE FROM N.A.
Complete proteome
                                                                 360
616
677
408
610
```

```
MCH_METTR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           NON_TER
                                                                                                                                                                                                                                                                                                                    09RPD4;
                                                                                                                                                                                                                                                                              RESULT 10
MCH_METIR
                                                                                                                                                                                                                                                                                                       DR
DR
DR
MW
MW
MW
MW
FT
FT
FT
FT
SO
                                                                                                                                                                                                                                        a
                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation. The European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce.or send an email to license@isb-sib.ch).
                                                          This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation-the European Bioinformatics Institute. There are no restrictions on its mase by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entitles requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@lsb-sib.ch).
                                                                                                                                                                                                                                                                              ö
                          SIMILARITY).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 TRANSFERRING ELECTRONS FROM SUCCINATE TO UBIQUINONE (COENZYME OF)
                                                                                                                                                                                                                                                                               Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                   01-NOV-1997 (Rel. 35, Created)
01-NOV-1997 (Rel. 35, Last sequence update)
15-SEP-2003 (Rel. 42, Last annotation update)
Putative succinate dehydrogenase cytochrome B subunit, mitochondrial
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       STRAIN-S288C / AB972;

BUNNEd-9169872;

BOWMAN S., Churcher C.M., Badcock K., Brown D., Chillingworth T.,

Connor R., Dedman K., Devlin K., Gentles S., Hamlin N., Hunt S.,

Jagels K., Lye G., Moule S., Odebl. C., Pearson D., Rajandream M.A.

Rice P., Skelton J., Walsh S., Whitchead S., Barrell B.G.;

The nucleotide sequence of Saccharomyces cerevisiae chromosome
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Saccharomyces cerevisiae (Baker's yeast).
Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
Saccharomycetales; Saccharomycetaceae; Saccharomyces.
                                                                                                                                                                                                                                                                              ó
           -1- PTM: PHOSPHORYLATED IN A CELL CYCLE-DEPENDENT MANNER (BY
                                                                                                                                                                                                                                                  71.4%; Score 35; DB 1; Length 574; 77.8%; Pred. No. 58;
                                                                                                                                                                                                                                                                              0; Indels
                                                                                                                                                                                   General Sponte; SPCC53.09c; -. pfam. PF04.058; DNA_POl_alpha. B; 1. DNA_POl_alpha. B; 1. DNA_POL_ation: Nuclear protein: Phosphorylation. SEQUENCE 574 AA: 64184 MW; 18C17FCFA7AA098A CRC64;
                                                                                                                                                                                                                                                                                                                                                                                              196 AA
                                                                                                                                                                                                                                                                              2; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                              PRT;
SUBCELLULAR LOCATION: Nuclear
                                                                                                                                                              EMBL; AL023704; CAA19261.1; -. PIR; T41395; T41395.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           EMBL; 249702; CAA89756.1; -.
                                                                                                                                                                                                                                                                Best Local Similarity 77.8
Matches 7; Conservative
                                                                                                                                                                                                                                                                                                                                                                                            STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       S54580; S54580.
S0004724; YMR118C.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Nature 387:90-93(1997).
                                                                                                                                                                                                                                                                                                                              205 VVVVGRIVV 213
                                                                                                                                                                                                                                                                                                      3 VVIVGRIVL 11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   KMR118C OR YM9718.17C
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            SEQUENCE FROM N.A
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     NCBI_TaxID-4932;
                                                                                                                                                                                                                                                                                                                                                                                            YM07_YEAST
Q04487;
                                                                                                                                                                                                                                                      Query Match
                                                                                                                                                                                                                                                                                                                                                                                  YMO7_YEAST
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         PIR;
SGD;
                                                                                                                                                                                                                                                                                                                                                                                            ò
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/
                                                                                                                                                                                                                                                                                                                                                                               ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Gaps
InterPro; IPROUDOUSED SABLEST: 1.

PROSITE: PSO11001: SDH_CYT: 1.

PROSITE: PSO11001: SDH_CYT_1: 1.

Hypothetical protein; Tricarboxylic acid cycle; Electron transport; Heme: Transmembrane: Mitochomicion; Transit peptide.

TRANSIT : 196 CYTOCHROME B SUBUNITAL.

CHAIN : 196 CYTOCHROME B SUBUNIT.

119 POTENTIAL.

TOWNWITAL.

TOWNSTATIAL.
                                                                                                                                                                                                                                                                                                                                                                               ;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
Methylocystaceae; Methylosinus.
NCBL_TaxID=426;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                28-FEB-2003 (Rel. 41, Created)
28-FEB-2003 (Rel. 41, Last sequence update)
28-FEB-2003 (Rel. 41, Last annotation update)
N(5), N(1b).nethenylterrahydromethanopterin cyclohydrolase
(EC 3.5.4.27) (Methenyl-H4MPT cyclohydrolase) (Fragment).
                                                                                                                                                                                                                                                                                                                           Score 34; DB 1; Length 196;
Pred. No. 34;
                                                                                                                                                                                                                                                                                                                                                                               2; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 third step.
-!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
-!- SIMILARITY: BELONGS TO THE MCH FAMILY.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 260 AA; 27622 MW; BD647C0DBF03C6A8 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      260 AA
                                                                                                                                                                                                                                                                                                                                                                            3; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            HAMAP; MF_00486; -; 1.
InterPro; IPR003209; Cyclohydrolase.
Pform: PF07289; MCH; 1.
Probom: PF011637; Cyclohydrolase; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Hydrolase; One-carbon metabolism.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              EMBL; AF162786; AAD56174.1; -.
                                                                                                                                                                                                                                                                                                                                   69.48;
                                                                                                                                                                                                                                                                                                                                                        54.58;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Methylosinus trichosporium
                                                                                                                                                                                                                                                                                                                              Query Match 69.4
Best Local Similarity 54.5
Matches 6; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                         179 GSIVLAGTLVL 189
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                             1 GSVVIVGRIVL 11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ; P94954; 1QLM.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SEQUENCE FROM N.A.
```

```
ENV_EIAV2
P22428;
                                    CARBOHYD
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CHAIN
TRANSMEM
TRANSMEM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        TRANSMEM
TRANSMEM
TRANSMEM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CARBOHYD
CARBOHYD
               CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CARBOHYD
               FT
                                                                                                                                                                                                                                                                                                                                                                                This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation-the Buropean Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for connection entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                           ö
                                                                                                                                                                                                                                                                                                               01-AUG-1991 (Rel. 19, Created)
01-AUG-1991 (Rel. 19, Last sequence update)
16-OCT-2001 (Rel. 40, Last annotation update)
ENV polyprotein precursor (Coat polyprotein) (Contains: Coat protein
6P90; Coat protein GP45).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         the Satis Interpreted to the Satis Interpreted
                                                           Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SEGUENCE FROM N.A.
PEDLINE-BRO72070; PubMed-2825406;
PAYDE S.L., Fang F.D., Liu C.P., Dhruva B.R., Rwambo P., Issel C.J.,
Montelaro R.C.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         "Antigenic variation and lentivirus persistence: variations in envelope gene sequences during EIAV infection resemble changes reported for sequential isolates of HIV.";
                                                           ;
;
     Score 34: DB 1; Length 260;
Pred. No. 43;
3: Mismatches 1; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                Equine infectious anemia virus (clone P3.2-1) (EIAV).
Viruses; Retroid viruscs; Retroviridae; Lentivirus.
NCBL_TaxID=11666;
                                                                                                                                                                                                                                                                         859 A.A.
     69.4%;
60.0%;
Query Match 69.4
Best Local Similarity 60.0
Matches 6; Conservative
                                                                                                                                                                                                                                                                         STANDARD;
                                                                                                                                                  121 GSVQVVGRVL 130
                                                                                                    1 GSVVIVGRIV 10
                                                                                                                                                                                                                        RESULT 11

DAY _ EIAVI

DAY _ EIAVI

DAY _ D1-202-1991

DT _ 01-302-1991

DE _ GP90; Coaff

RA _ HONCELTAXID-

RA _ HONCELATO FR

CC _ C _ EERO FR

CC _ C _ EERO FR

CC _ C _ ENTINE FR

FT _ CARBOHYD

FT 
                                                                                                                                                a
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss institute of Bioinformatics and the EMEL outsitution. The between the Date institution institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                              ö
                                                                                                                                                                                                                                                                                                                                                      01-AUG-1991 (Rel. 19, Created)
01-AUG-1991 (Rel. 19, Last sequence update)
16-OCT-2001 (Rel. 40, Last annotation update)
ENV polyprotein precursor (Coat polyprotein) (Contains: Coat protein
GP90; Coat protein GP45).
                                                                                                                                 Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                SEQUENCE FROM N.A.
BEDLINE-88072070; Pubmed-2825406;
Payne S.L., Fang F.D., Liu C.P., Dhruva B.R., Rwambo P., Issel C.J.,
Montelaro R.C.;
557 557 N-LINKED (GLCNAC. . .) (POTENTIAL).
752 752 N-LINKED (GLCNAC. . .) (POTENTIAL).
859 AA; 97140 MW; F4AOC071396DA867 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           (POTENTIAL)
(POTENTIAL)
(POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (POTENTIAL)
(POTENTIAL)
(POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        "Antigenic variation and lentivirus persistence: variations in envelope gene sequences during EIAV infection resemble changes reported for sequential isolates of HIV."; virology 161:321-331(1987).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            EMBL; M18386; AAA66408.1; -.
PIR; B34027; VCL_E2.
InterPro; IPR001027; Gp45_EIAV.
InterPro; IPR001021; Gp90_EIAV.
Pfam; PF01045; EIAV_GF90; 1.
Pfam; PF00971; EIAV_GF90; 1.
Coat protein; Glycoprotein; Polyprotein; Transmembrane; Signal.
SIGNAL.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              (POTENTIAL)
                                                                                                                              ö
                                                                             Query Match 69.4%; Score 34; DB 1; Length 859; Best Local Similarity 77.8%; Pred. No. 1.3e+02; Matches 7; Conservative 1; Mismatches 1; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Equine infectious anemia virus (clone P3.2-2) (EIAV). Viruses; Retroid viruses; Retroviridae; Lentivirus. NCBI_TaxID=11667;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         COAT PROTEIN GP90.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               POTENTIAL.

N-LINKED (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ENV POLYPROTEIN.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (GLCNAC
                                                                                                                                                                                                                                                                                                                     859 A.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       POTENTIAL. POTENTIAL.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          N-LINKED
N-LINKED
                                                                                                                                                                                                                                                                                                                  STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     | | |:|||||
| BO1 GLVIIVGRI 809
                                                                                                                                                                  1 GSVVIVGRI 9
```

```
SEÓUENCE FROM N.A.
MEDLINE-87236196; PubMed-3035786;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ž.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 irology 158:300-312(1987).
          214
233
244
346
346
368
368
368
406
411
411
422
483
550
557
97066
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Query Match
Best Local Similarity 7/.0.
Best Local 7; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              REVISIONS TO N-TERMINUS.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    | |:|||||
| 801 GLVIIVGRI 809
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1 GSVVIVGRI 9
                 proviral DNA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Tronick S.R.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ENV_EIAV9
P11306;
                                            CARBOHYD
CARBOHYD
                                                                                                                                        CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                        CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                    CARBOHYD
SEQUENCE
                                                                                                                                                                                                                                                                                                         CARBOHYD
                                                                                                                CARBOHYD
                                                                                                                                                                                                             CARBOHYD
                                                                                                                                                                                                                                                                                                                                    CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                          CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                     CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            RESULT 14
ENV_EIAV9
                    HERE THE LEAD OF A STREET OF THE STREET OF T
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    This SWISS-PROT entry is copyright, It is produced through a collaboration between the Swiss Institute of Bloinformatics and the EMBL outstation-the European Bloinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@lsb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                               ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  01-AUG-1991 (Rel. 19, Created)
01-AUG-1991 (Rel. 19, Last sequence update)
16-OCT-2001 (Rel. 40, Last annotation update)
ENV polyprotein precursor (Coat polyprotein) [Contains: Coat protein GP90; Coat protein GP45].
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         MEDIJNE-88072070; PubMed-2825406; Pagna S.L., Fang F.D., Liu C.P., Dhruva B.R., Rwambo P., Issel C.J., Pagne S.L., Fang F.D., Liu C.P., Dhruva B.R., Rwambo P., Issel C.J., Montelaro R.C., "Antigenic variation and lentivirus persistence: variations in envelope gene sequences during EIAV infection resemble changes virology 161:321-331(1987).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Gaps
      (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ( POTENTIAL) ( POT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Polyprotein; Transmembrane; Signal
                                                                                                                                                                                                                                                                                                                                                                                           Score 34; DB 1; Length 859;
Pred. No. 1.3e-02;
1; Mismatches 1; Indels
N-LINKED (GLCNAC. . . ) (POTE N-LINKED (GLCNAC. . ) (POTE N-LINKED N
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Equine infectious anemia virus (clone P3.2-3) (EIAV).
Viruses: Retroid viruses; Retroviridae; Lentivirus.
NCBL_TaxID=11668;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   POTENTIAL.
ENV POLYPROTEIN.
COAT PROTEIN GP90.
COAT PROTEIN GP45.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           N-LINKED (GLCNAC.
N-LINKED (GLCNAC.
N-LINKED (GLCNAC.
N-LINKED (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 POTENTIAL. POTENTIAL.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                POTENTIAL. POTENTIAL.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              EMBL, M18387; AAA66409.1; -- PIR, CSA077; VCLLB3. InterPro. 197001027; Gp45_EIAV. InterPro; 197001361; Gp90_EIAV. Pfam.; PF001045; EIAV_GP90; EIAV_GP90; EIAV_GP90; I. Pfam.; PF001971; EIAV_GP90; I. Coat protein; Glycoprotein; Polyf
                                                                                                                                                                                                                                                                                                                                    3
                                                                                                                                                                                                                                                                                                                                                                                              69.4%;
77.8%;
                                                                                                                                                                                                                                                                                                                                    97188
                                                                                                                                                                                                                                                                                                                                                                                              Query Match 69.4
Best Local Similarity 77.5
Matches 7; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   22
859
859
859
859
863
636
835
1112
1141
186
      368
3999
3999
4411
4483
4483
4490
550
7527
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  | |:|||||
| 801 GLVIIVGRI 809
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1 GSVVIVGRI 9
368
399
411
421
422
483
483
490
550
550
557
752
752
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ENV_EIAV3
P22429;
                                                                    CARBOHYD
CARBOHYD
CARBOHYD
                                                                                                                                                                    CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                  CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              TRANSMEM
TRANSMEM
TRANSMEM
   CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 TRANSMEM
TRANSMEM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SIGNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CHAIN
CHAIN
CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          RESULT 13
ENV_EIAV3
   FT
FT
FT
FT
FT
FT
FT
SO
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              The state of the s
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the Buropean Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    01-JUL-1989 (Rel. 11, Created)
01-FEB-1996 (Rel. 33, Last sequence update)
16-CCT-2001 (Rel. 40, Last annotation update)
ENV polyprotein preursor (Coat polyprotein) [Contains: Coat protein
GP90; Coat protein GP45].
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                N-LINKED (GLCNAC...) (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Rawakami T., Sherman L., Dahlberg J., Gazit A., Yaniv A.,
Tronick S.R., Aaronson S.A.;
"Nucleotide sequence analysis of equine infectious anemia virus
                                                                                                                                                                                                                                                                                                                                                                                                                         69.4%; Score 34; DB 1; Length 859; 77.8%; Pred. No. 1.3e+02; ive 1; Mismatches 1; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Submitted (NOV-1987) to the EMBL/GenBank/DDBJ databases.
                                                                                                                                                                                                                                                                                                                                                                         982A9F5A1AD8FA4D CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Equine infectious anemia virus (clone 1369) (EIAV).
Viruses; Retroid viruses; Retroviridae; Lentivirus.
NCBL_TaxID=11670;
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the Burpoena Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ·,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SEQUENCE FROM N.A.
MEDLINE-92292330; PubMed-1318398;
Petry S.T., Flaherty M.T., Kelley M.J., Clabough D.L., Tronick S.R.,
Coggins L., Whetter L., Lengel C.R., Fuller F.;
"The surface envelope protein gene region of equine infectious anemia
virus is not an important determinant of tropism in vitro.";
J. Virol. 66:4085-4097(1992).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    BNV_EIAVC STANDARD; PRT; 859 AA. p3.2641; b3.2641; b3.2641; created) b1-0cT-1993 (Rel. 27, Last sequence update) b1-0cT-1993 (Rel. 40, Last annotation update) b16-0cT-2001 (Rel. 40, Last annotation update) coat protein GP90; Coat protein GP45].
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Gaps
                                                                                                                      (POTENTIAL)
(POTENTIAL)
(POTENTIAL)
(POTENTIAL)
(POTENTIAL)
(POTENTIAL)
(POTENTIAL)
(POTENTIAL)
(POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                        (POTENTIAL).
(POTENTIAL).
(POTENTIAL).
(POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                               (POTENTIAL)
                                                                      POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Polyprotein; Transmembrane; Signal. POTENTIAL.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Length 859
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1; Indels
                         N-LINKED (GLCNAC...) (POTINED 
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Equine infectious anemia virus (clone CL22) (EIAV).
Viruses; Retroid viruses; Retroviridae; Lentivirus.
NCBI_TaxID=31675,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ENV POLYPROTEIN.
COAT PROTEIN GP90.
COAT PROTEIN GP45.
POTENTIAL.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     69.4%; Score 34; DB 1; I
77.8%; Pred. No. 1.3e+02;
iive 1; Mismatches 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                EMBL: M87581; AAA43005.1; --
PIR: C41991: VCLJ22.
InterPro: IPR001027; Gp45_E1AV.
InterPro: IPR0010361; Gp90_E1AV.
Pfam: PF01045; EIAV_GP45; 1.
Coat protefn: Glycoprotein; PolygsIGNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  97113 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              001 GLVIIVGRI 809
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1 GSVVIVGRI 9
Best Local Similarity
Matches 7; Conserv
                                                                                            CARBOHYD
CAR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CHAIN
CHAIN
TRANSMEM
TRANSMEM
TRANSMEM
CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CHAIN
```

```
ö
                                                                                                                                                                                                                                                       Gaps
                                                                                                                                                             (POTENTIAL).
                                                             (POTENTIAL).
(POTENTIAL).
(POTENTIAL).
                                                                                                  (POTENTIAL)
(POTENTIAL)
(POTENTIAL)
(POTENTIAL)
                                                                                                                                         (POTENTIAL)
                               (POTENTIAL)
                                                                                                                                                                                         N-LINKED (GLCNAC. ..) (POTENTIAL.)
N-LINKED (GLCNAC. ..) (POTENTIAL.)
23E020E80DF334FA CRC64;
                                                                                                                                                                                                                                                      ·.
                                                                                                                                                                                                                                  Length 859;
                                                                                                                                                                                                                                 Score 34; DB 1; Length 859
Pred. No. 1.3e+02;
1; Mismatches 1; Indels
                             N-LINKED (GLCNAC.
                       POTENTIAL
                                                                                                                                                                                                                                                                                                                        Search completed: August 30, 2003, 19:13:51 Job time : 1.54479 secs
                                                                                                                                                                                                              M.
                                                                                                                                                                                                                                  69.4%;
77.8%;
636
835
807
807
811
1112
1112
1112
1148
1148
1244
340
360
368
369
369
369
401
550
557
7557
                                                                                                                                                                                                                       Query Match
Best Local Similarity 77.0
Tricology 77.0
Tricology 77.0
                                                                                                                                                                                                                                                                                   1 GSVVIVGRI 9
                                                                                                                                                             CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
SEQUENCE
                             CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
                                                                               CARBOHYD
CARBOHYD
                                                                                                   CARBOHYD
          TRANSMEM
TRANSMEM
                                                                                                                     CARBOHYD
                                                                                                                                                    CARBOHYD
                                                                                                                                 CARBOHYD
                                                                                                                                          CARBOHYD
   q
```

ð

```
hepatitis
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            MEDLINE-9370154; Pubmed-10438800;

A MEDLINE-9370154; Pubmed-10438800;

R Coch J.O., Bartenschlager R.;

*Modulation of hepetitis C virus NS5A hyperphosphorylation by nonstructural proteins NS3. NS4A, and NS4B.*;

J. Virol. 73:7138-7146(1999).

-1. SUBDINIT: HEW VIRIAL OF THIS VIRUS IS A NUCLECCAPSID COVERED BY A LIPOPROTEIN ENVELOPE CONSISTS OF TWO PROTEINS:

PROTEIN M AND GLYCOPPORTEN E. THE NUCLECCAPSID IS A COMPLEX OF PROTEIN M AND GLYCOPPORTEN E. THE NUCLECCAPSID IS A COMPLEX OF PROTEIN M AND GLYCOPPORTEN E. THE NUCLEOCAPSID IS A COMPLEX OF REMEL; AJ238800; CAB53095.1; -.

REMEL; AJ238800; CAB53095.1; -.

REMEL; AJ238800; CAB53095.1; -.

RICEPPO; IPRO01410; DEAD.

RICEPPO; IPRO01521; HCV_copeid.

RICEPPO; IPRO02521; HCV_core.

RICEPPO; IPRO02531; HCV_NS3.

RICEPPO; IPRO01409; HCV_NS3.

RICEPPO; IPRO0145; HCV_NS3.

RICEPPO; IPRO0145; HCV_NS3.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Hepatitis C virus type 1b.
Viruses; ssRNA positive strand viruses, no DNA stage: Flaviviridae;
Hepacivirus.
  068210
068227
068227
068209
068216
068221
068221
068221
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
06
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    STRAIN-RB;
Bartenschlager R.;
Submitted (MAY-1999) to the EMBL/GenBank/DDBJ databases.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Last sequence update)
Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                PRT; 3010 AA
                                                                                                                                                                                                                                                                                                                                                                                                    ALIGNMENTS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Created)
                                      068207
068216
068218
068228
                                                                           068221
068205
068215
068232
068232
                                                                                                                                                                Q68238
Q68237
Q68217
                                                                                                                                                                                                     068230
081579
081577
081575
081575
                                                                                                                                                                                                                                                                    081584
081574
081578
081583
081581
081573
068344
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        01-MAY 2000 (TrEMBLrel, 13, 01-MAY-2000 (TrEMBLrel, 13, 01-MAR-2003 (TrEMBLrel, 23, Genome polyprotein,
 PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             [2]
SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   NCBI_TaxID-31647;
 STRAIN-RB;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                Q9QP06;
Q9QP06;
                                                                                                                                                                                                                                                                                                                                                                                                                                      RESULT 1
Q9QP06
099p61 heptilis c 081760 heptilis c 081760 heptilis c 091fff mucosal dis 03914 heptilis c 03985 heptilis c 039x5 heptilis c 068209 heptilis c 068218 heptilis c 06824 heptilis c
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    09qp06 hepatitis c
09qp61 hepatitis c
081760 hepatitis c
                                                                                 August 30, 2003, 19:00:22; Search time 2.09905 Seconds (without alignments) 1352.314 Million cell updates/sec
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Pred. No. 1s the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Description
                                                                                                                                                                                                                                                         830525
          GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.
                                                                                                                                                                                                                                                     Total number of hits satisfying chosen parameters:
                                                                                                                                                                                                                              830525 seqs, 258052604 residues
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SUMMARIES
                                                                                                                                                                                                                                                                                                                    Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries
                                                              using sw model
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    090P06
090P61
081760
090TE5
039914
039895
09978P5
068235
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               068244
068206
068242
068240
068240
                                                                                                                                                                                        BLOSUM62
Gapop 10.0 , Gapext 0.5
                                                                                                                                                                                                                                                                                                                                                                                                        sp_fundi: *
sp_fundi: *
sp_human: *
sp_mammal: *
sp_mhc: *
sp_phage: *
sp_phage: *
sp_phage: *
sp_phage: *
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          sp_vertebrate:*
sp_unclassified:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               sp_bacteriap:*
                                                                                                                                                                                                                                                                                                                                                                      SPTREMBL_23:*
: sp_archea:*
: sp_bacteria:*
                                                                                                                                                                                                                                                                              Minimum DB seq length: 0
Maximum DB seq length: 2000000000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   sp_rvirus:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      sp_rodent:*
                                                                                                                                    US-09-965-594-26
49
                                                                                                                                                               1 GSVVIVGRIVE 11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              sp_archeap:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Sp_virus:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            DB
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    3010
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             3010
3010
4040
                                                             protein search,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1000.0
1000.0
1000.0
1000.0
1000.0
98.0
98.0
98.0
98.0
98.0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Query
                                                                                                                                                                                         Scoring table:
                                                                                                                                        Title:
Perfect score:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Score
                                                                protein
                                                                                                                                                                                                                               Searched:
                                                                                                                                                                  Sequence:
                                                                                                                                                                                                                                                                                                                                                                        Database
                                                                                     Run on:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             647978976749
647449
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Result
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              2
```

```
IPR001490;
                                                                                                                          InterPro; InterP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      InterPro;
                                                                                   InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            InterPro;
                                                                                                        InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         081760
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     RESULT 3
Q81760
                                                                                   ő
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ...v_N84b.
...v_N82868; HCV_N85a.
...v_N012868; HCV_N85a.
...v_N012868; HCV_N85a.
...v_N012868; HCV_RARP.
DR InterPro; IPR007095; RNA_POL_DS_PS.
DR Pfam; PP01543; HCV_corb; 1.
DR Pfam; PP01543; HCV_corb; 1.
DR Pfam; PP01543; HCV_corb; 1.
DR Pfam; PP01560; HCV_N81; 1.
DR Pfam; PP01560; HCV_N81; 1.
DR Pfam; PP01066; HCV_N82; 1.
DR Pfam; PP01006; HCV_N84; 1.
DR Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Genome polyprotein.
Hepatitis C virus.
Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
Hepatvirus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    8 NSSA PHOSPHOPROTEIN.
0 NS5B RNA DEPENDANT RNA POLYMERASE.
326999 MW; AS70BB980DD64634 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        100.0%; Score 49; DB 12; Length 3010; 100.0%; Pred. No. 6.6;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   0; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 01-MAY-2000 (TrEMBLrel. 13, Created)
01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
01-WAR-2003 (TrEMBLrel. 23, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               PRT; 3010 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   0; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      InterPro; IPR001410; DEAD.
InterPro; IPR00252; HCV_capsid.
InterPro; IPR002521; HCV_core.
InterPro; IPR002519; HCV_env.
InterPro; IPR002519; HCV_NSI.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     11; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             PRELIMINARY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1 GSVVIVGRIVL 11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              810
1027
1658
1712
1973
2419
3010 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  NCBI_TaxID+11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           090P61
090P61;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  RESULT 2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  q
```

```
PROSITE: PS50507; RDRP_POSITIVE; 1.
PROSITE; PS50521; RDRP_VIRAL; 1.
ATP-binding; Coat protein; Envelope protein; Glycoprotein; Helicase; Hydrolase; Nonstructural protein; Polyprotein;
RNA-directed RNA polymearase; Transferase; Transmembrane.
SEQUENCE 3010 AA; 327068 MW; 9105F69483DD5BBA CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Hepatitis C virus.
Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
Hepacivirus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ;
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       100.0%; Score 49; DB 12; Length 3010; 100.0%; Pred. No. 6.6; tive 0; Mismatches 0; Indels 0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Last sequence update)
Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   PRT; 3010 A.A.
                                InterPro; IPR001400; HCV_NG4b.
InterPro; IPR001400; HCV_NG4b.
InterPro; IPR001668; HCV_NG5b.
InterPro; IPR001650; Helicase_C.
InterPro; IPR001650; Helicase_C.
InterPro; IPR001050; RNA_pol_DS_PS.
InterPro; IPR001094; RNA_pol_DS_PS.
InterPro; IPR001094; RNA_pol_DS_PS.
InterPro; IPR001094; RNA_pol_DS_PS.
InterPro; IPR001094; RNA_pol_DS_VIT.
Pfam; PF01543; HCV_core; I.
Pfam; PF01559; HCV_Core; I.
Pfam; PF01560; HCV_NS1; I.
Pfam; PF01560; HCV_NS2; I.
Pfam; PF01001; HCV_NS3; I.
Pfam; PF01001; HCV_NS3; I.
Pfam; PF01001; HCV_NS3; I.
Pfam; PF0101; HCV_NS3; I.
Pfam; PF0101; HCV_NS3; I.
Pfam; PF0101; HCV_NS3; I.
Pfam; PF001909; VITAL_RARP; I.
Pr0Dom; PD186062; HCV_NS1; I.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Created)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                HCV_capsid.
HCV_core.
HCV_env.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     01-NOV-1996 (TrEMBLrel. 01, 01-NOV-1996 (TrEMBLrel. 01, 01-MAR-2003 (TrEMBLrel. 23,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1678 GSVVIVGRIVL 1688
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Local Similarity 100.
nes 11; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1 GSVVIVGRIVL 11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   IPR004109;
IPR000745;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Genome polyprotein.
Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                NCBI_TaxID-11103;
```

HCV\_NS4b

ö

```
Lai V.C., Zhong W., Skelton A., Ingravallo P., Vassilev V.,
Donis R.O., Hong Z., Lau J.Y.;
"Generation and characterization of a hepatitis C virus NS3 protease-
dependent bowine viral diarrhea virus.";
J. Virol. 74:6339-6347(2000).
                                                                                                                                                                                                                                                                                                                                                                                                    PROSITE; PRO0487; DEXDC; 1.
PROSITE; PS00190; CYTOCHROME_C; 2.
PROSITE; PS50507; RDRP_POSITIVE; 1.
ATP-binding; Coat protein; Envelope protein; Glycoprotein; Helicase; Hydrolase; Monstructural protein; Polyprotein; BNA-directed RNA polymerase; Transmembrane.
SEQUENCE 3010 AA; 326986 MM; 573C0F3C55B3F3F4 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Mucosal disease virus.
Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             100.0%; Score 49; DB 12; Length 3010; 100.0%; Pred. No. 6.6;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   0; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      J. Virol. 74:6339-6347(2000).
[2] SEQUENCE FROM N.A.
Lai V.C.H., Hong Z.;
Submitted (MAY-2000) to the EMBL/GenBank/DDBJ databases.
EMBL; AF266278; AAF82566.1; -.
HSSP; P26663; 1JXP.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         01-OCT-2000 (TrEMBLrel. 15, Created)
01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 PRT; 4040 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               0; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Interpro; IPR000280; CDvir_endptseP80.
Interpro; IPR0001410; DEAD.
Interpro; IPR001410; DEAD.
Interpro; IPR001410; HCV_NS3.
Interpro; IPR00150; HCV_NSRP.
Interpro; IPR00150; HG11case_C.
Interpro; IPR001005; Myb_DNA_binding.
Interpro; IPR001005; RNA_pOl_DS_PS.
Interpro; IPR001095; RNA_pOl_DS_PS.
Interpro; IPR007094; RNA_pOl_DS_PS.
Interpro; IPR007094; RNA_pOl_PSvir.
Pfam; PF00291; HCJ_RGRP; I.
Pfam; PF00291; HCJ_RGRP; I.
Pfam; PF00398; VIral_RGRP; I.
Pfam; PF00398; VIral_RGRP; I.
SMART; SM00487; DEXDC; I.
SMART; SM00480; HELICC; I.
SMART; SM00480; HELICC; I.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                [1]
SEQUENCE FROM N.A.
MEDLINE-20323484; Pubmed-10864644;
                                  pfam; pF01542; HCV_core; 1. pfam; pF01542; HCV_core; 1. pfam; pF01560; HCV_NS1; 1. pfam; pF01560; HCV_NS2; 1. pfam; pF01500; HCV_NS3; 1. pfam; pF01000; HCV_NS4; 1. pfam; pF01001; HCV_NS4s; 1. pfam; pF01506; HCV_NS5s; 1. pfam; pF001506; HCV_NS5s; 1. pfam; pF001506; HCV_NS5s; 1. pfam; pF001506; HCV_NS5s; 1. pfam; pF00998; Viral_RGRP; 1. propression; p1186062; HCV_NS1; 1. cvapm; pr000407; propression; p1186062; HCV_NS1; 1. cvapm; pr000407; pr000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1678 GSVVIVGRIVL 1688
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            11; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1 GSWIVGRIVL 11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Genome polyprotein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             NCBI_TaxID-11099;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               MEROPS; S31.001;
      PF01543;
PF01542;
PF01539;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Q91FH8
Q91FH8;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            RESULT 5
091FH8
             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 DOLLAR REPAREMENT OF THE DESCRIPTION OF THE PRESENT OF THE PRESENT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Prodom; PD186062; HCV_NS1; 1.
SMART; SMOM6487; DEXDC: 1.
PROSITE; PS50507; RDRP_VORAL; 1.
Coat protein; Bysolope protein; Glycoprotein; Nonstructural protein;
Polyprotein; RNP-directed RNA polymerase; Transferse; Transmembrane.
SEQUENCE 3010 AA; 326955 MW; EA7D306A4BA2E224 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Viruses; s9RNA positive strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Takahashi K., Iwata K., Matsumoto M., Matsumoto H., Nakao K.,
Hatahara T., Ohta Y., Kanai K., Maruo H., Baba K., Hijikata M.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           0;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         100.0%; Score 49; DB 12; Length 3010; 100.0%; Pred. No. 6.6:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           0; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     01-MAR-2001 (TrEMBLrel. 16, Created)
01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             PRT; 3010 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           0; Mismatches
                        InterPro; IPR002166; HCV_RGRP.
InterPro; IPR007095; RNA_PO1_DS_PS.
InterPro; IPR007094; RNA_PO1_DS_PS.
InterPro; IPR007094; RNA_PO1_DS_PS.
InterPro; IPR007094; RNA_PO1_PSVIr.
Pfam; PP01542; HCV_caps_id; I.
Pfam; PP01542; HCV_caps_id; I.
Pfam; PP01560; HCV_NS2; I.
Pfam; PP01006; HCV_NS3; I.
Pfam; PP01006; HCV_NS4; I.
Pfam; PP01001; HCV_NS4b; I.
Pfam; PP01001; HCV_NS4b; I.
Pfam; PP01001; HCV_NS5a; I.
Pfam; PP01001; HCV_NS5a; I.
Pfam; PP01001; HCV_NS5a; I.
Pfam; PP01000; HCV_NS5a; I.
Pfam; PP01000; HCV_NS5a; I.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         RNA_pol_DS_PS.
RNA_pol_PSvir.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      HCV_RdRP.
Helicase_C.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             InterPro; IPR002522; HCV_capsid.
InterPro; IPR002521; HCV_core.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  HCV_NS2.
HCV_NS3.
HCV_NS4a.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       HCV_NS4b.
InterPro; IPR002868; HCV_NS5a.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        InterPro; IPR002519; HCV_env.
InterPro; IPR002531; HCV_NS1.
InterPro; IPR002518; HCV_NS2.
InterPro; IPR004109; HCV_NS3.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1678 GSVVIVGRIVL 1688
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Local Similarity 100.
Les 11; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1 GSVVIVGRIVL 11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 IPR000745;
IPR001490;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            IPR007094;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         IPR001650;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Genome polyprotein.
Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          IPR002166;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SEQUENCE FROM N.A. STRAIN-HCVT145;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         NCBI_TaxID-11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Hepacivirus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Mishiro S.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            InterPro;
InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Q9DTE5;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                RESULT 4

OUTES

1000TES

1000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Matches
   g
```

ö

Gaps

; 0

```
J. Gen. Virol. 76:211-215(1995)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Best Local Similarity 90.9
Matches 10; Conservative
                                                                                                                                                                                                                                                             Local Similarity 90.9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    20 GSVVIVGRIIL 30
                                                                                                                                                                                                                                                                                                                                                     1 GSVVIVGRIVL 11
                                                                                                                                                                                                                                                                                                                                                                                       8 GSVVIVGRIIL 18
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1 GSVVIVGRIVL 11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              NCBI_TaxID=11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              NCBI_TaxID=11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  102
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (Fragment)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Query Match
                                                                                                                                                                                                                                          Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        09PXP5:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    068209;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      068209
                                                                                                                                                                                                                                                                                              Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              RESULT 8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  RESULT 9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Q9PXP5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ó
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            PER SERVICE SE
         STRRETS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         엄
                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ÷
                                                                                                                                                                                                                                                                ·;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                "Sequence analysis of hepatitis C virus variants producing discrepant results with two different genotyping assays.";
J. Med. Virol. 53:237-244(1997).
EMBL: AF007519; AB602970.2;
InterPro: IPR00745; HCV.MS4a.
Pfam: PF01006; HCV_NS4a: 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              "Sequence analysis of Hepatitis C virus variants producing discrepant results with two different genotyping assays.";
                     PROSTIE; PS50521; RNRP VIRAL; 1.
PROSITE; PS00531; RNASE_I2_2; 1.
ATF-Didding; Helicase; Hydrolase; Nonstructural protein; Polyprotein; SRNA-directed RNA polymerase; Transferase.
SEQUENCE 4040 AA; 453073 MW; ADE87791D055B9DC CRC64;
                                                                                                                                                                                                                                                                   Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Hepatitis C virus.
Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
Hepacivirus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Viruses; ssRNA positive-strand viruses, no DNA stage: Flaviviridae;
Hepacivirus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Prescott L.E., Berger A., Pawlotsky J.-M., Conjeevaram P., Pike I., Simmonds P.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     STRAIN-FD_1;
BDLINE-198032593; PubMed-9365889;
Prescott L.E., Berger A., Pawlotsky J.M., Conjeevaram P., Pike I.,
Simmonds P.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     .;
0
                                                                                                                                                                                                                                                                ö
                                                                                                                                                                                             Query Match
Best Local Similarity 100.0%; Pred. No. 8.8;
Matches 11; Conservative 0; Mismatches 0; Indels 0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Score 48; DB 12; Length 88;
Pred. No. 0.3;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     0; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Created)
Last sequence update)
Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Last sequence update)
Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                89 A.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                88 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        01-JNN-1998 (TEMBLEA). 05, Created)
01-JNN-2002 (TEMBLEA). 21, Last Sequent
01-0CT-2002 (TEMBLEA). 22, Last annotat
Non-structural protein 4a/b (Fragment).
NS4A/B.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             01-JAN-1998 (TrEMBLrel. 05, Created)
01-DEC-2001 (TrEMBLrel. 19, Last sequen
01-MAR-2002 (TrEMBLrel. 20, Last annota
Non-structural protein 4a/b (Fragment).
PS50507; RDRP_POSITIVE; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           98.08;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Local Similarity 90.5
nes 10; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             PRELIMINARY;
                                                                                                                                                                                                                                                                                                                          1 GSVVIVGRIVL 11
                                                                                                                                                                                                                                                                                                                                                        10 GSVVIVGRIVE 20
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1 GSWIVGRIVI, 11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               NCBI_TaxID-11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   NCBI_TaxID-11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           NON_TER
NON_TER
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         039895
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          039914
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    NS4A/B
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Best Loc
Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                   RESULT 6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      RESULT 7
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              039914
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SO WWW SO SO WW SO WW SO WW SO W SO W
```

```
÷
                                                                                                                                                                                                                                                                  ;;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SEQUENCE FROM N.A.
MEDLINE-94245087; PubMed-7514558;
Tanaka T., Tsukiyama-Kohara K., Yamaguchi K., Yagi S., Tanaka S.,
Hanseqawa A., Ohta Y., Hattori N., Kohara M.;
"Significance of specific antibody assay for genotyping of hepatitis C
                                                                                                                                                                                                                                                                     Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
Hepacivirus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
Hepacivirus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           MEDLINE=95146953; PubMed=7844535;
Greene W.K., Cheong M.K., Ng V., Yap K.W.;
"Prevalence of hepatitis C virus sequence variants in South-East
Asia.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ö
                                                                                                                                                                                                                                                                  .;
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   01-MAY-2000 (TrEMBLrel. 13, Created)
01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
01-JUN-2002 (TrEMBLrel. 1). Last annotation update)
NON-structural protein NS4-GROUP II HCV-specific antigen C14-1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Length 102;
                                                                                                                                                                                                         Score 48; DB 12; Length 89;
Pred. No. 0.3;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 0; Indels
                                                                                                                                                                                                                                                                0; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SEQUENCE 102 AA; 11419 MW; 08124C19CF367F06 CRC64;
J. Med. Virol. 0:0-0(1997).

EMBL; AF007500; AAB62951.1; -.

Interp.: IPROGUA5; HCV_NS4a.

Pfam; PF01006; HCV_NS4a.

Pfam; PF01006; HCV_NS4a.

NON.TER 89 89

SEQUENCE 89 AA; 9748 MW; 581BBBC8A3EA8B5C CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  01-NOV-1996 (TrEMBLrel. 01, Created)
01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
Nonstructural protein (Fragment).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             98.0%; Score 48; DB 12;
90.9%; Pred. No. 0.35;
Live 1; Mismatches 0.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   138 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    102 AA
                                                                                                                                                                                                                                                          1; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Hepatology 19:1347-1353(1994).
InterPro; IPR000745; HCV_NS4a.
Pfam; PF01006; HCV_NS4a: 1.
                                                                                                                                                                                                           98.08;
90.98;
```

S

ö

Gaps

;

```
Hepatitis C virus.
Viruses; sRRNA positive-strand viruses, no DNA stage; Flaviviridae;
Hepacivirus.
NCB1_TaxID=11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
Hepacivirus.
NCBL_TaxID=11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           STRAIN-1b;
MEDLINE-55146953; PubMed-7844535;
Greene W.K., Cheong M.K., Ng V., Yap K.W.;
Greene W.E. cheong M.K., Ng V., Yap K.W.;
Spravalence of hepatitis C virus sequence variants in South-East Asia.";
J. Gen. Virol. 76:211-215(1995).
         "Prevalence of hepatitis C virus sequence variants in South-East
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       98.0%; Score 48; DB 12; Length 138; 90.9%; Pred. No. 0.47; tive 1; Mismatches 0; Indels
                                                                                                                                                                                                                                                                                                                                  Query Match 98.0%; Score 48; DB 12; Length 138; Best Local Similarity 90.9%; Pred. No. 0.47; Matches 10; Conservative 1; Mismatches 0; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 EMBL; U14280; AAC53969.1; -...
HSSP; P27958; IHEI.
InterPro; IPR000745; HCV_NS4a.
Pfam, PF01006; HCV_NS4a; 1.
NON_TER 138 138
SEQUENCE 138 AA; 15118 MM; B7F7EB2733770408 CRC64;
                                                                                                                                                                                                                                            01-NOV-1996 (TrEMBLrel. 01, Created)
01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
Nonstructural protein (Fragment).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              01-NOV-1996 (TrEMBLrel. 01, Created)
01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
Nonstructural protein (Fragment).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              138 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         138 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 PRT;
                                                       EMBL, Uirol. 76:211-215(1995).
EMBL, U14254; AAC53943.1;
HSSP; P27958; IHEI.
Interpro; IPR000745; HCV_NS4a.
RAM; PF01006; HCV_NS4a; I.
NON_TER 138 138
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Local Similarity 90.9 nes 10; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                         1 GSVVIVGRIVL 11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1 GSVVIVGRIVL 11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                     Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             068244;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  068206;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              068244
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         968206
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Rest Loc
Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               RESULT 13
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       RESULT 12
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           068244
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SO DRANKER REPRESENTATION OF THE PROPERTY OF T
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   a
         RT RE DR DR DR DR DR DR DR SO SO
                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Q
                                                                                                                                                                                                                                                                                                      ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Gaps
                                                                                                                                                                                                                                                                                                         Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
Hepacivirus.
NCBI_TaxID=11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae; Hepacivirus.
WCBL_TaxID=11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Greene W.K., Cheong M.K., Ng V., Yap K.W.;
"Prevalence of hepatitis C virus sequence variants in South-East
Asla.";
                                                                                                                                                                                                                                                                                                      .;
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ö
                                                                                                                                                                                                                                      Query Match 98.0%; Score 48; DB 12; Length 138; Best Local Similarity 90.9%; Prod. No. 0.47; Matches 10; Conservative 1; Mismatches 0; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Query Match 98.0%; Score 48; DB 12; Length 138; Best Local Similarity 90.9%; Pred. No. 0.47; Matches 10; Conservative 1; Mismatches 0; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Interpro; IPR000745; HCV_NS4a.
Pfam: PF01006; HCV_NS4a; 1.
NON_TER 1 1 1
NON_TER 138
SEQUENCE 138 Aa: 15104 MW; 585DC5A627D0F3E3 CRC64;
                                                                                                                 NON_TER 1 1 1 NON_TER 138 138 SEQUENCE 138 AA; 15149 WW; DBAE62A0FE9E2D57 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         01-NOV-1996 (TIEMBLIEL 01, Created)
01-NOV-1996 (TIEMBLIEL 01, Last sequence update)
01-DEC-2001 (TIEMBLIEL 19, Last annotation update)
Nonstructural protein (Fragment).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     01-NOV-1996 (TIEMBLrel. 01, Created)
01-NOV-1996 (TIEMBLrel. 01, Last sequence update)
01-DEC-2001 (TIEMBLrel. 19, Last annotation update)
Nonstructural protein (Fragment).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     138 A.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                138 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    STRAIN-1D;
MEDLINE-95146953; PubMed-7844535;
Greene W.K., Cheong M.K., Ng V., Yap K.W.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     STRAIN-1D;
MEDLINE-95146953; PubMed-7844535;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               J. Gen. Virol. 76:211-215(1995).
EMBL: U14271; AAC53960.1; -.
HSSP; P27958; 1HEI.
EMBL; U14245; AACS3934.1; -.
HSSP; P27958; 1HEI.
INTERPRO; IPRO00745; HCV_NS4a.
Pfam; PF01006; HCV_NS4a; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                             1 GSVVIVGRIVL 11
                                                                                                                                                                                                                                                                                                                                                                                               1 GSVVIVGRIVL 11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    068235;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       068218
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  068235
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              RESULT 10
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   RESULT 11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         HANGE ON THE REPORT OF THE THE REPORT OF THE
   DR DR DR SO FT SO SO
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                à
```

ö

Gaps

ö

```
NCBI_TaxID-11103;
     ô
                                                                                                                                                                                                                                                                                                     a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ö
                                                                                                                                                                                                             ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Gaps
C STRAIN-1b;

X HEDLINE-95146953; PubMed-7844535;

A Greene W. K., Cheong M.K., Ng V., Yap K.W.;

T "Prevalence of hepatitis C virus sequence variants in South-East II Asia.";

J. Gen. Virol. 76:211-215(1995).

R EMBL: 014242, AAC53931.1; -.

R HSSP: P27958; JHEI.

Interpro: IPR000745; HCV_NS4a.

R Pfam: PF01006; HCV_NS4a.

R NON_TER 138 A3: 15117 MW; 5FBC51A1B74DE13E CRC64;
                                                                                                                                                                                                             Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Hepatitis C virus.
Viruses: ssRNA positive-strand viruses, no DNA stage; FlavIvirldae;
Hepacivirus.
                                                                                                                                                                                                                                                                                                                                                                                                                    Hepatitis C virus.
Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
Hepacivirus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Greene W.K., Cheong M.K., Ng V., Yap K.W.; 
 "Prevalence of hepatitis C virus sequence variants in South-East
                                                                                                                                                                                                             .;
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            98.0%; Score 48; DB 12; Length 138; 90.9%; Pred, No. 0.47; U.ve 1; Mismatches 0; Indels
                                                                                                                                                                                98.0%; Score 48; DB 12; Length 138; 90.9%; Pred. No. 0.47; tive 1; Mismatches 0; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SEQUENCE 138 AA; 15281 MW; CD5B5B3834C6070D CRC64;
                                                                                                                                                                                                                                                                                                                                                      01-NOV-1996 (TTEMBLrel. 01, Created)
01-NOV-1996 (TTEMBLrel. 01, Last sequence update)
01-DEC-2001 (TTEMBLrel. 19, Last annotation update)
Nonstructural protein (Fragment).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 01-NOV-1996 (TrEMBLrel. 01, Created)
01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
Nonstructural protein (Frägment).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          PRT; 138 AA
                                                                                                                                                                                                                                                                                                                               138 AA
                                                                                                                                                                                                                                                                                                                               PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                STRAIN-1b;
MEDLINE-95146953; Pubmed-7844535;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             J. Gen. Virol. 76:211-215(1995).
EMBL; U14278; AAC53967.1; ...
HSSP; P27958; 1HEI.
Interpro; IPR000745; HCV_NS4a.
Pfam: PF01006; HCV_NS4a; 1.
                                                                                                                                                                                   Query Match 98.0
Best Local Similarity 90.9
Matches 10; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Local Similarity 90.9 es 10; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1 GSVVIVGRIVL 11
                                                                                                                                                                                                                                      1 GSVVIVGRIVL 11
                                                                                                                                                                                                                                                     52 GSVVIVGRIIL 62
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         138
                                                                                                                                                                                                                                                                                                                                                                                                                                                                         [1]
SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                          NCBI_TaxID-11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Q68240
Q68240;
                                                                                                                                                                                                                                                                                                                               068242
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             RESULT 15
068240
1D 068244
AC 068244
DT 01-NO'
DT 01-NO'
DT 01-DB
DE NONST
GN NS4.
CS HEPAT
OC HEPAT
                                                                                                                                                                                                                                                                                                     RESULT 14
Q68242
    RX
RX
RX
RX
RX
RX
DR
DR
DR
SQ
SQ
                                                                                                                                                                                                                                                                                                                                ò
```

```
RN | 13
RP SEQUENCE FROM N.A.
RC STRAIN-lb;
RX MEDLINE-95146953; Pubmed-7844535;
RA Greene W.K., Cheong M.K., Ng V., Yap K.W.;
RT Asia.";
RT Prevalence of hepatitis C virus sequence variants in South-East
RT PASIA.";
RL J. Gen. Virol. 76:211-215(1995).
REMBL: 014276; AAC53651: -.
DR HSSP; P27958; 1HEI.
DR HSSP; P27958; 1HEI.
DR HSSP; P27958; 1HEI.
DR HSSP; P27958; 1HEI.
DR HSSP; P1970006; HCV_NS4a: 1.
FT NON_TER 138 138
SQ SEQUENCE 138 AA. 15115 MW; 6A042677B354CA7A CRC64;
CHORY MAtch
SQ SEQUENCE 138 AA. 15115 MW; 6A042677B354CA7A CRC64;

QUETY MAtch
SQ SEQUENCE 138 AA. 15115 MW; 6A042677B354CA7A CRC64;

QUETY MAtch
SQ SEQUENCE 138 AA. 15115 MW; 6A042677B354CA7A CRC64;

QO I GSVVIVGRIVL 11

| | | | | | | | | | | | | |
DD 52 GSVVIVGRIVL 11

Search completed: August 30, 2003, 19:18:23
JOB time : 4.09905 secs
```

```
Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ALIGNMENTS
                                                                                                                                                SUMMARIES
                                                                                                                                                                                      E06198
E06394
A91965
AR031992
AR207294
HCJ238800
AB049092
                                                                                                                                                                                                                                              HPCCGENOM
HPCRNA
                                                                                                                                                                                                                                                                                                                                                                                                                    HCU14241
HCU14242
HCU14243
HCU14244
HCU14245
HCU14245
HCU14245
                                                                                                                                                                                                                                                              AF176573
AR179057
AF268278
                                                                                                                                                                                                                                                                                                                                                                             E09291
AF007500
                                                                                                                                                                                                                                                                                                                                                             AF007519
                                                                                                                                                                                                                                                                                     ARI45217
ARI45217
ARI45197
ARI45213
ARI45213
ARI45247
ARI45249
AX481515
                                                                                                                                                                                                                                                                                                                                                                     AR037528
                                                                                                                                                                                                                                                                                                                                                                                                      E11062
E03389
                                                                                                                                                                                                                                                                                                                                                      E09290
                                                                                                                                                                                                                                                                                                                                                                                              E11063
                                                                                         em_htgo_mus:*
em_htgo_other:*
                em_htg_inv:*
em_htg_other:*
                                                                                em_htgo_hum:*
        em_htg_hum:*
                                 em_htg_mus:
                                        em_htg_pln:*
                                                 em_htg_rod:*
                                                        em_htg_mam:*
em_htg_vrt:*
                                                                                                                                                                        DB
                                                                                                                                                               Query
Match Length
                                                                                                                                                                                                                                                                                                                                    161
189
189
266
267
267
                                                                                                                                                                                                                               9033
                                                                                                                                                                                                      Score
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             RESULT 1
                                                                                                                                                                                                                                                                                       Result
                                                                                                                                                                                                                                                                    -WODEL-frame-_p2n.model -DEV=xlp
-Q-Cqqq2_1_XCBFQ_spool_VGSQ965594_runat_29082003_151919_28110/app_query.fasta_1.2872
-Q-Cqqq21_XCBFQ_spool_VGSQ965594_runat_29082003_151919_28110/app_query.fasta_1.2872
-Q-Cqqq2_1_XCBFQ_spool_VGSQ965594_runat_29082003_151919_2810
-UNITS-bits -START=1 -END=-1 -MATRIX-blosum62 -TRANS-human40.cd1 -LiST=45
-DCGALIGN=200 -THR_SCORE-FCT -THR_MX=100 -THR_MN=0 -ALIGN=15 -MODE-LOCAL
-OUTFHT-pto -NORM-ext -HEAPSIZE-500 -MINITSH-0 -MAXLEN-200000000
-USER-USG9965594_GCGN_1 1_14686_Grunat_29082003_151919_28310 -NCPU-6 -ICPU-3
-NO_MARP -LARGEQUERY NEG_SCORES-0 -WAIT -DSPBLOCK=100 -LONGLOG
-FGAPOT-120 -WARN TIMEOUT-30 -THREADS-1 -XAAPOP-10 -XGAPEXT=0.5 -FGAPOP-6
-FGAPEXT=7 -YGAPOP-10 -YGAPEXT=0.5 -DELEXT=7
                                                      August 30, 2003, 19:18:33; Search time 142.976 Seconds (without alignments) 3147.423 Million cell updates/sec
      GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.
                                       nucleic search, using frame_plus_p2n model
                                                                                                                                                                      2888711 segs, 20454813386 residues
                                                                                                                                                                                      hits satisfying chosen parameters:
                                                                                                                                                                                                                             Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries
                                                                                                                               0.5
                                                                                                                              Xgapop 10.0 , Xgapext
Ygapop 10.0 , Ygapext
Fgapop 6.0 , Fgapext
Delop 6.0 , Delext
                                                                                                                                                                                                      Minimum DB seq length: 0
Maximum DB seq length: 2000000000
                                                                                      US-09-965-594-26
                                                                                                       1 GSVVIVGRIVL 11
                                                                                                                                                                                                                                                                                                                                                                   9b_ba:*
9b_htg:*
9b_in:*
9b_on:*
9b_pat:*
9b_pat:*
9b_pr:*
9b_pr:*
9b_rro:*
9b_rro:*
9b_vro:*
9b_vro:*
9b_vro:*
9b_vro:*
9b_vro:*
9b_vro:*
9b_vro:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          em_in:
em_mu:*
em_om:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   em_or:*
em_ov:*
em_pat:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           em_ph:*
em_pl:*
em_ro:*
em_sts:*
                                                                                                                                                                                                                                                                                                                                                             GenEmbl: *
                                                                                                                                                                                                                                                              Command line parameters:
                                                                                                                        BLOSUM62
                                                                                                                                                                                                                                                                                                                                                                                                                                            ot
                                                                                               Perfect score:
                                                                                                                        Scoring table:
                                        OM protein -
                                                                                                                                                                                      Total number
                                                                                                                                                                                                                                                                                                                                                             Database :
                                                                                                       Sequence:
                                                                                                                                                                       Searched:
                                                       Run on:
                                                                                        ritle:
```

AR011992 Sequence AR207294 Sequence AJ238800 Hepatitis AB049092 Hepatitis L02836 Hepatitis C D10934 Hepatitis C AR176573 Hepatitis C AR17657 Sequence

E06198 cDNA encodi

Description

AF268278 Pestiviru

AR145217 Sequence AR145221 Sequence

AR14521 Sequence AR145197 Sequence AR145213 Sequence AR145247 Sequence AR145249 Sequence AR81515 Sequence AR807527 Sequence E09290 DNA encodin AF007519 Hepatitis

E09291 DNA encodin AF007500 Hepatitis E11062 DNA sequenc E11062 DNA sequenc E11062 DNA sequenc E11062 DNA sequenc U14241 Hepatitis C U14243 Hepatitis C U14245 Hepatitis C U14245 Hepatitis C U14251 Hepatitis C U14251 Hepatitis C U14254 Hepatitis C U14264 Hepatitis C

```
Hepocivitus.

Seki,M., Honda,Y., Takahashi,K., Murakami,T., Teranishi,Y. and Hayashi,N., Honda,Y., Takahashi,K., Murakami,T., Teranishi,Y. and Hayashi,N., Honda,Y., Takahashi,K., Murakami,T., Teranishi,Y. and Hayashi,N. ERB DENDE PROM HEPATITIS C VIRUS, POLYPEPTIDE CODED BY THE SAME AND ITS PRODUCTION PROMISE HILLS CORP.

MITSUBISHI KASEI CORP.

OS (Hepatitis C virus)
PN JP 1994000086-A/38
PD 11-3M-1994
PF 07-CCT-1992 JP 1992293734
PR 07-CCT-1992 JP 91P 287008, 16-DEC-1991 JP 91P 332329, PR 20-ARN-1992 JP 92P 99957
PI SEKI MAKOTO, HONDA YOSHIKAZU, TAKAHASHI KAZUNOBU, PI HURAKAMI TOMORO,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   circular PAT 22-JAN-2000
           JP 199400086-A/38.
Hepaitis C virus
Hepaitis C virus
Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                       CizN15/85,C12P21/02//A61K39/29,(C12P21/02,C12R1:91),C07K99:00;
strandedness: Double;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 9. .608
/product-'the peptides reacting specifically and
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                immunochemically with the serum of hepatitis
                                                                                                                                                                                                                                                                                                                                        TERANISHI YUTAKA, HAYASHI NORIO
C12N15/51,C07K7/06,C07K7/08,C07K7/10,C07K13/00,C12N5/10,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           618
0
0
0
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Location/Qualifiers
1. 618
/organiam="Hepatitis C virus"
/mol_type="genomic RNA"
/db_xref="taxon:11103"
a 184 c 179 g 120 t
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ONA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1 GlySerValValIleValGlyArgIleValLeu 11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1 (bases I to 7475)
Attwood, M.R. and Hurst, D.N.
ATHIVINAL PEPTIDE DERIVATIVES
PATENT: WO 9822496-A I 28-MAY-1998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Length:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                A91965 7475 bp
Sequence 1 from Patent W09822496.
A91965 A91965.1 G1:6740811
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               patient'.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     US-09-965-594-26 (1-11) x E06394 (1-618)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           179 9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   type C
                                                                                                                                                                                                                                                                                                                                                                                                                                                    *source: clone=N13-1;
                                                                                                                                                                                                                                                                                                                                                                                                                     topology: Linear;
anti-sense: No;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         0.238
49.00
100.00%
100.00%
100.00%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   mat_peptide
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              unidentified unclassified.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 unidentified
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Percent Similarity:
Best Local Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ď
                                                                                                                                                                                                                                                                                                                                        PI TERAN
C12N1
C12N1
C12N1
C12N1
CC C12N1
CC STIAN
CC antl-
CC antl-
CC *sour
FH Key
FT mat_F
FT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               135
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Query Match:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  source
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ORGANISM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               AUTHORS
TITLE
JOURNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               BASE COUNT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ..
9
                                                  ORGANISM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                REFERENCE
                                                                                               REFERENCE
AUTHORS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   RESULT 3
A91965
VERSION
KEYWORDS
SOURCE
                                                                                                                                                                             JOURNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 FEATURES
                                                                                                                                               TITLE
                                                                                                                                                                                                           COMMENT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ORIGIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               PAT 29-SEP-1997
                                                                                                                              Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
Hepacivirus.
1 (bases 1 to 618)
Seki,M., Honda,Y., Takahashi,K., Murakami,T., Teranishi,Y. and
                  PAT 29-SEP-1997
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   C12R1:19), (C12N5/10, C12R1:91), (C12P21/02, C12R1:19), (C12P21/02,
                                                                                                                                                                                                                                                                                                                                                                                                                     PI TERANISHI YUTAKA, HAYASHI NORIO
C CLINIS/S1,0007K7/06,C07K7/08,C07K7/10,C07K13/00,C07K15/12, PC
C12N1/21,C12N5/10,
PC
C12N15/11,C12N15/70,C12N15/85,C12P21/02//A61K39/00,A61K39/29,
                                                                                                                                                                                                                                                                                                                                        172794, PR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             9. .608
/product-'the peptides reacting specifically
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              immunochemically with the serum of hepatitis
                                                                                                                                                                                     Hayashi,N.
GENE OR DNA FRAGMENT DERIVED FROM HEPATITIS C VIRUS, POLYPEPTIDE CODED BY THE SAME AND ITS PRODUCTION
PACENT: JP 1994000085-A 38 11-JAN-1994;
MITSUBISHI RASEI CORP
                                                                                                                                                                                                                                                                                                                                                       332329, PR
                                                                                                                                                                                                                                                                                                                                                                             PI SEKI MANOTO, HONDA YOSHIKAZU, TAKAHASHI KAZUNOBU, PI
MURAKAMI TOMOKO,
                                                                                                                                                                                                                                                                      PD 11-JAN-1994

PF 11-JAN-1994 JP 1992194497

PR 11-JUN-1991 JP 91P 139268, 12-JUL-1891 JP 91P

07-CCT-1991 JP 91P 287008, 16-DEC-1991 JP 91P 3322

PI SEKI MAKOTO ".....
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CDNA encoding genes derived from hepatitis C virus. E06394
                18 bp RNA linear encoding genes derived from hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    618
00
00
00
00
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Length:
Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              /___618
/organism="Hepatitis C virus"
/mol_type="genomic RNN"
/db_xref="taxon:11103"
a 184 c 179 g 120 t
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Gaps:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             type C
    patient'.
Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 US-09-965-594-26 (1-11) x E06198 (1-618)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  strandedness: Double;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 *source: clone=N13-1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               and
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   topology: Linear;
anti-sense: No;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  0.238
49.00
100.008
100.008
                                                                 E06198.1 GI:2174385
                                                                                JP 1994000085-A/38
                                                                                               Hepatitis C virus
Hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                mat_peptide
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       (C12N1/21,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Score:
Percent Similarity:
Best Local Similarity:
Query Match:
DB:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Ø
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         135
                                                  E06198
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             DEFINITION
ACCESSION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            source
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       BASE COUNT
ORIGIN
                LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ..
Q
                                                                                                                                                            REFERENCE
AUTHORS
                                                                                                                                                                                                                                           JOURNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               RESULT 2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             FEATURES
                                                                                                                                                                                                           TITLE
                                                                                                                                                                                                                                                                           COMMENT
 E06198
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ð
```

ᠬ

```
Location/Qualifiers
1. .7475
                       /organism="unknown"
i 1916 c 2039 g
                                    2039 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                      2 (bases 1 to 9033)
Bartenschlager, R.
Direct Submission
                                                                             3.53
49.00
100.00%
100.00%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1. ,9033
                                                                                                     Percent Similarity:
Best Local Similarity:
                                   1853 a
                                                                                                                                                                                                                                                                                                                                                                                                                                               99370154
                                                                                                                                                                                                                                                                                                                                                                                                                                                           10438800
                                                                      Alignment Scores:
                                                                                                                          Query Match:
DB:
                                                                                                                                                                                                                                                                 DEFINITION
ACCESSION
VERSION
KEYWORDS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       source
                                                                                                                                                                                                                                                                                                                                             SOURCE
ORGANISM
                                  BASE COUNT
ORIGIN
                                                                               Pred. No.:
                                                                                                                                                                                                                                RESULT 6
HCJ238800
                                                                                                                                                                                                                                                                                                                                                                                                     AUTHORS
TITLE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                      REFERENCE
AUTHORS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             TITLE
JOURNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                     JOURNAL
MEDLINE
                                                                                                                                                                                                                                                                                                                                                                                          REFERENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                             PUBMED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CDS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           FEATURES
  FEATURES
                                                                                                                                                                                                           셤
                                                                                                                                                                                                                                                                                       PAT 29-SEP-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          PAT 20-JUN-2002
                                                                                                                                                                                                                                                                                                                                                                          1 (bases 1 to 7475)
Attwood, M. Richard., Hurst, D. Nigel., Jones, P. Stephen.,
Kay, P. Brittain., Raynham. T. Michhael. and Wilson, F. Xavier.
Peptidyl inbitors of viral proteases
Patent: US 5866684-A 1 02-FEB-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1 (bases 1 to 7475)
Attwood, M. Richard., Hurst, D. Nigel., Jones, P. Stephen.,
Kay, P. Brittain., Raynham, T. Michael. and Wilson, F. Xavier.
Antiviral medicaments
Patent: US 6372883-A 1 16-APR-2002;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          linear
                                                                                                              7475
11
0
0
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  7475
111
0
0
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Conservative:
Mismatches:
Indels:
                                                                                                                                     Conservative:
Mismatches:
                                                                                                                                                                                                                                DNA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1 GlySerValValIleValGlyArglleValLeu 11
                                                                                                                                                                                                                     1 GlySerValValIleValGlyArgIleValLeu 11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ب
                                                                   1667 t
                                                                                                               Length:
Matches:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Length:
Matches:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1667
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Sequence 1 from patent US 6372883.
AR207294 GI:21506162
                                                                                                                                                             Indels:
                                                                                                                                                                                                                                                                                  Sequence 1 from patent US 5866684. AR031992. AR031992.1 GI:5946281
                              /organism="unidentified"
/mol_type="genomic DNA"
/db_xref="taxon:32644"
i 1916 c 2039 g 166
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 US-09-965-594-26 (1-11) x AR031992 (1-7475)
                                                                                                                                                                                              US-09-965-594-26 (1-11) x A91965 (1-7475)
                                                                                                                                                                                                                                                                                                                                                                                                                                                       /organism-"unknown"
1916 c 2039 q
          Location/Qualifiers
1. .7475
                                                                                                                                                                                                                                                                                                                                                                                                                                      Location/Qualifiers
LA ROCHE (CH)
                                                                                                              3.53
49.00
100.00%
100.00%
100.00%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  3.53
49.00
100.00%
100.00%
100.00%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Unclassified.
                                                                                                                                                                                                                                                                                                                                                                   Unclassified.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Percent Similarity:
Best Local Similarity:
Query Match:
                                                                                                                                                  Best Local Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1853 a
                                                                      ď
HOFFMANN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Unknown.
                                                                                                                                                                                                                                                                                                                                            Unknown.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Unknown
                                                                   1853
                                                                                                                                                                                                                                                                                                                                                         Unknown
                                                                                                                                     Percent Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Alignment Scores:
                                                                                                      Alignment Scores:
                                                                                                                                                         Query Match:
DB:
                                                                                                                                                                                                                                                                                                                                KEYWORDS
SOURCE
ORGANISM
                                                                                                                                                                                                                                                                                                                                                                                                                                                 source
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   DEFINITION
                                                                  BASE COUNT
ORIGIN
                                                                                                                                                                                                                                                              AR031992
LOCUS
                                                                                                                                                                                                                                                                                               DEFINITION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                     BASE COUNT
ORIGIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ORGANISM
                                                                                                                  Pred. No.:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Pred. No.:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ACCESSION
VERSION
KEYWORDS
SOURCE
                                                                                                                                                                                                                                                                                                           ACCESSION
VERSION
                                                                                                                                                                                                                                                                                                                                                                                                              TITLE
JOURNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    TITLE
JOURNAL
                                                                                                                                                                                                                                                                                                                                                                             REFERENCE
                                                                                                                                                                                                                                                                                                                                                                                         AUTHORS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    REFERENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                AUTHORS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   RESULT 5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               AR207294
                                                                                                                                                                                                                                                                                                                                                                                                                                      PEATURES
          FEATURES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
```

```
HCJ238800 9033 bp RNA linear VRL 18-AUG-1999
Hepatitis C virus type 1b complete genome, isolate NCI.
AJ28800
AJ238800.1 GI:5748510
complete genome: core protein; glycoprotein E1; glycoprotein E2;
NS2 proteinase; NS3 proteinase/Helicase; NS3/4A proteinase
cofactor; NS4b protein; NS5A phosphoprotein; NS5B RNA dependant RNA
polymerase; p7 peptide; p0lyprotein; NS5B RNA dependant RNA
Hepatitis C virus type 1b
Hepatitis C virus type 1b
Viruses; SSRNA positive-strand viruses, no DNA stage; Flaviviridae;
Hepacivirus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               GVRATEKTSERSOPRGEROPITKAROPEGRAWAOPGYPWPLYCNEGLGWAGWILSPRG
SRPSWGPTDPRRRSRNLGKVIDTLTCGFADLMGYLPLVGAPLGGAARALAHGVRVLED
GVNYATGNLPGCSFSIFLLALLSCLTPASATEWTNSGCYPTTWTNDCSNASITYTEAAD
WHITPGCVPCVTENTSRCWVALITTLAANASPPTTTIRRHYDLLVGAAALCSANY
VGDLCGSVFLYAQLFRFSPRHETVODCMCSIYPGHYTGHRAMDMANMWSPTAALVY
SQLLRIPQAVVDMVAGAHWGYLAGLAYYSMVGNWARVLIVMLLFAGVDGGTYYTGGTM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    AKNTIGITSLFSPGSSQKIQLVNTNGSWHINRTALNCNDSLNTGFLAALFTVHKFNSS
GCPERWASCSPIDAFAQGWGPITYNESHSSDQRPYCHTYAPRCGITVAAQVGGPVYC
FYSPVVVGTTDRFGVPTFSRGENETDVLLLNNTRPDQGNWFGCTWNSTGFTKTGG
PPCNIGGIGSKTLLFGPTDCFRKHPEATYTKGGSGPWILPRCLVHYPYRLWHYPCTVNF
TIFKVRMYVGGVEHRLEAACNWTRGERCNLEDRDRSELSPLLLSTTEWQVLPCSFTTL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   /codon_start-1
/product-*polyprotein*
/protein_id="CAB53095.1*
/db_xref-*GI:5148511*
/translation=*MSTNPRPQRKTRRNTNRRPQDVKFPGGGQIVGGVYLLPRRGPRL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Submitted (10-MAY-1999) Bartenschlager R., Institute for Virology, Johannes Gutenberg - University Mainz, Obere Zahlbacher Strasse 67, 55131 Mainz, GERMANY Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               β
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Koch,J.O. and Bartenschlager,R.
andlation of hepatits C virus NS5A hyperphosphorylation
nonstructural proteins NS3, NS4A, and NS4B
j. virol. 73 (9), 7138-7146 (1999)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         /organism="Hepatitis C virus type 1b"
                           Matches:
Conservative:
Mismatches:
                                                                                                                                                                                                                                                                                     Indels:
Length:
                                                                                                                                                                  Gaps:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           /virion
/mol_type="genomic RNA"
/strain="RB"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   /isolate="NC1"
/db_xref="taxon:31647"
/country="Germany"
                                                                                                                                                                                                                            US-09-965-594-26 (1-11) x AR207294 (1-7475)
```

```
RIGITARY PROBERES DEGRALOUS TO THE LITTLY PARA STANDARD READER RE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              KVVILDSFEDLRAEEDEREVSVPAEILRRSRKFPRAMPIWARPDYNPPLLESWKDPDY
VPPVVHGCPLPPAKAPPIPPPRKRTVVLSESTVSSALAELATKTFGSSESSAADSGT
ATAPPDQPSDDGDAGSDVGSYSSMPPLEGEPGDPDLSDGSWSTVSEEAGEDVVCCSMS
                                                                                                                                                                                                           DSRGSLLSPRPVSYLKGSSGGPLLCPSGHAVGIFRAAVCTRGVAKAVDFVPESMETT
MRSPVFTDNSSPPAVPQTFQVAHLHAPTGSGKSTKVPAAYAAGGYKVI.VLNPSVAATL
SFGAYMSKAHGIDPNIRTGVRTTTTGAPITYSTYGKFLADGGCSGGAYDIIICDECHS
                                                                                                                                                                                                                                                                                                                                   TDSTTILGIGTVLDQAETAGARLVVLATATPPGSVTVPHPNIEEVALSSTGEVPFYGK
AIPIETIKGGRHLIFCHSKKKCDELAAKLSGLGLNAVAYYRGLDVSVIPTSGDVIVVA
TDALATGFTGDFDSVIDCNTCVTQTVDFSLDPTFTIETTTVPQDAVSRSQRRGRTGRG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          YTWTGALITPCAAEETKLPINALSNSLLRHHNLYYATTSRSASLROKKYTFDELQYLD
DHYRDVLKEMKAKASTVKAKLLSVEEACKLTPPHSARSKFGYGAKDVRNLSSKAVSHI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              VVSTLPQANGSSYGFQTSPGGRVEELVNTWKSKKCPMGFAYDTRCFDSTVTREBDINV
EESIYQCCDLAPEARQAIRSLTERLYIGGPLTNSKGQNCGYRRCRASGVLTTSCGNTL
TCYLKAAAACRAAKLODCTMLVCGDDLVVICESAGTQEDEASLRAFTEAMTRYSAPPG
LOVWIPPLNVRGGRDAVILLTCAIHPELIFTITKILLAILGPLMVLQAGITKVPYYVR
                                       AHGLIRACMIVRKVAGGHYVQMALMKLAALTGTYVYDHI,TPLRDWAHAGLRDLAVAVE
                                                                                                                            AYSQQTRGLLGCIITSLTGRDRNQVEGEVQMVSTATQSFLATCVNGVCWTVYHGAGSK
TLAGPKGPITOMYTNVDQDLVGWQAPPGARSLTPCTCGSSDLYLVTRHADVIPVRRRG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          RSVWKDLLEDTETPIDTTIMAKNEVFCVQPEKGGRKPARLIVFPDLGVRVCEKMALYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          DPPRPEYDLELITSCSSNVSVAHDASGKRYYYLLRDPTTPLARAAWETARHTPVNSWL
GNIHYAPTLWARMILMTHFPSILLAQEQLEKALDCQIYGACYSIEPLDLPQIIQRLH
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              GLSAFSLHSYSPGEINRVASCLRKLGVPPLRVWRHRARSVRARLLSQGGRAATCĞKYL
FNWAVRTKLKLTPIPAASQLDLSSWFVAGYSGGDTYRSLSRARPRWFMWCLLLLSVGV
                                                                                    PVVFSDMETKVITWGPDTAACGDIILGLPVSARRGREIHLGPADSLEGQGWRLLAHII
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             'product" NS3 proteinase/helicase
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    /product-"glycoprotein El"
1150. 2238
/gene-"E2"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              /product-"glycoprotein E2"
2239. .2427
/gene-"p7"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    /product="NS2 proteinase"
3079. .4971
/gene="NS3"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        /product-"core protein"
574, 1149
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     /product-"NS4B protein"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        /product="p7 peptide"
2428. 3078
/gene="NS2"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   /gene="E2"
2239. ...
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      /gene="E1"
/gene="E1"
1150.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         'gene="core"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         J/4 .1149
/gene="E1"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 . .573
'gene="core"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             /gene="NS2"
3079. .4971
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     5134. .5916
/gene="NS4B
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         3078
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 /gene="p7"
2428. .3078
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              GIYLLPNR"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1. .573
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      nat_peptide
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  mat_peptide
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             mat_peptide
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         mat_peptide
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     mat_peptide
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              mat_peptide
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          mat_peptide
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          gene
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          gene
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         gene
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              gene
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             gene
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     gene
```

```
Hepatitis C virus (HCV) genotype 1b sequences from fifteen patients with hepatocellular carcinoma: the 'progression score' revisited Hepatol. Res. 20 (2), 161-171 (2001)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         VRL 22-AUG-2002
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Hepatitis C virus
Hepatitis C virus
Viruses; SSRNA positive-strand viruses, no DNA stage; Flaviviridae;
Hepacivirus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Takahashi,K., Iwata,K., Matsumoto,M., Matsumoto,H., Nakao,K.,
Hatahara,T., Ohta,Y., Kanai,K., Maruo,H., Baba,K., Hijikata,M. and
Mishiro,S.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Direct Submission
Submitted (19-SEP-2000) Shunji Mishiro, Toshiba General Hospital,
Submitted (19-SEP-2000) Shunji Mishiro, Toshiba General Hospital,
Department of Medical Sciences; 6-3-22 Higashi Oh-i, Shinagawa-ku,
Tokyo 140-8522, Japan (E-mail:shunji.mishiro@po.toshiba.co.jp,
Tel:81-3-3764-8981, Pars:81-3-3764-8992)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Hepatitis C virus gene for polyprotein, complete cds, isolate:HCVT145.
AB049092.1 GI:11559450
                                                                                                                                                                                                                                                                 /product="NS5B RNA dependant RNA polymerase"
2731 c 2574 g 1903 t
4972. 5134
/gene="NS4A"
/product="NS3/4A proteinase cofactor"
5134. .5916
/gene="NS4B"
5917. .7256
                                                                                                                                                                                                                                                                                                                                               9033
11
0
0
0
                                                                                                                                                                                                                                                                                                                                               Length:
Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                     5917. 7256

/gene-"NS5A"

5917. 7256

/gene-"NS5A"

/product-"NS5A phosphoprotein"

7257. 9030

/gene-"NS5B"

/gene-"NS5B"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1 GlySerValValileValGlyArgileValLeu 11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           /organism="Hepatitis C virus"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       /specific_nost=human"
/db_xref="taxon:11103"
/tissue_type='serum'
/note='HCV-genotype: 1b'
<1. 310
                                                                                                                                                                                                                                                                                                                                                                                                                                Gaps:
                                                                                                                                                                                                                                                                                                                                                                                                                                                            US-09-965-594-26 (1-11) x HCJ238800 (1-9033)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          /mol_type="genomic RNA" /isolate="HCVT145"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         /note="partial"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            (bases 1 to 9344)
                                                                                                                                                                                                                                                                                                                                               4.33
49.00
100.00%
100.00%
100.00%
                                                                                                                                                                                                                                                                                                                                                                                                Best Local Similarity:
                                                                                                                                                                                                                                                                                  1825 a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Mishiro, S
                                                                                                                                                                                                                                                                                                                                                                               Percent Similarity:
                                                                                                                                                                                                                                                                                                                                 Alignment Scores:
                                                                                                                                                       mat_peptide
                                                                                                                                                                                                                                     mat_peptide
                                          mat_peptide
                                                                                                                                                                                                                                                                                                                                                                                                               Query Match:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     DEFINITION
                                                                                                                                                                                                                                                                                  BASE COUNT
                                                                                                                                                                                                                                                                                                                                                  Pred. No.:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       5'UTR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ORGANISM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          TITLE
JOURNAL
                                                                                                                         gene
                                                                                                                                                                                                      gene
             gene
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ACCESSION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  REFERENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 AUTHORS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             JOURNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            REFERENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            AUTHORS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       RESULT 7
AB049092
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      VERSION
KEYWORDS
SOURCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            FEATURES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               TITLE
                                                                                                                                                                                                                                                                                                   ORIGIN
```

CDS

```
HPDALILIS C virus RNA, complete genome sequence.

Hepatitis C virus RNA, complete genome sequence.

D10934.1 GI:471116
C; E; NS1/E2; NS2; NS3; NS4; NS5; polyprotein precursor.

Hepatitis C virus
Viruses; SSRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                         VRL 02-AUG-1993
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Wang,Y., Okamoto,H., Tsuda,F., Nagayama,R., Tao,Q.M. and Mishiro,S. Prevalence, genotypes, and an isolate (HC-C2) of hepatitis C virus in Chinese patients with liver disease J. Med. Virol. 40 (3), 254-260 (1993)
                                                                                                                                                                                                                                                                                                                                                                       Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
Hepacivirus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Original source text: Hepatitis C virus (individual_isolate China, strain Hebel) (library: NC1-18) cDNA to genomic RNA.
Location/Qualifiers
1. 9400
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         These data kindly submitted in computer readable form by: Hiroaki
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      from mainland of
                                                                                                                                      lincar
                                                                                                                                                                                                                                                                                                                                                                                                                                   l (bases 1 to 9400)
Bi.S.-1., Bai.X.-H., Margolis,H.S. and Liu,C.-B.
Complete hepatitis C virus genome cDNA isolated
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         9400
111
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Conservative:
                                                                                                                              HPCCGENOM 9400 bp ss-RNA LO2836 China virus complete genome. L02836.1 G1:329739 complete genome; viral genome. Hepatitis C virus Hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          5342 GGCAGCGTGGTCATTGTGGGCAGGATCGTCCTG 5374
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        /organism-"Hepatitis C virus"
/mol_type="genomic RNA"
/strain-"Hebei"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          2650 g 2034 t
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Length:
Matches:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Gaps:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              US-09-965-594-26 (1-11) x HPCCGENOM (1-9400)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                /isolate="China"
/db_xref="taxon:11103"
/tissue_lib="NC1-18"
355. .9384
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                (bases 1 to 9456)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Jichi Medical School
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         4.52
49.00
100.00%
100.00%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Immunology Division
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      2796 c
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Unpublished (1992)
Original source tea
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Kawachi-gun
Tochigi 329-04
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Percent Similarity:
Best Local Similarity:
Query Match:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ø
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1920
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      8394876
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Okamoto
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               source
                                                                                                                              LOCUS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      BASE COUNT
ORIGIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Pred. No.:
                                                                                                                                                                                                 ACCESSION
VERSION
KEYWORDS
                                                                                                                                                                                                                                                                                                                                ORGANISM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                REFERENCE
AUTHORS
TITLE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               JOURNAL
MEDLINE
PUBMED
                                                               RESULT 8
HPCCGENOM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                      AUTHORS
TITLE
                                                                                                                                                                                                                                                                                                                                                                                                                                          REFERENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      JOURNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            RESULT 9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        FEATURES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     COMMENT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               HPCRNA
                                                                                                                                                                                                                                                                                                    SOURCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Score:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                IAQUEPALEHLIAPARAYCAGANUSYU YIKWEYYU KIKELULLE LILLAPARAYCAGAMMUL
IAQUEPALHAMUUNUNUYILIAALIGAHGILESTIN YIKHEYLILE LILLAPPRAYAMDEBAASGGAVEVGISLUYYUKAETAKLIWMLQALITRAEHL
LULALPPRAYAMDEBAASGGAVEVGILALITISPHYKVELAKLIWMLQYLITRAEHL
LOWHYPEPUNKGGERAMILLATGTYYNHLITPIENDWAHAGLIBUPAYA
AQGIVRACHI, VYYONYULMKILAALIGTYYNHLITPIENDWAHAGLIBUPAYA
AQGIVRACHI, YYONYOYULMKILAALIGTYYNHLITPIENDWAHAGLIBUPAYA
ASQCTRGLLGGITTSELTGRDKNQVEGEVQVYSTATGSELAYCHYNGYCWYYYTHADYIPYRRG
DSRGSLLSPRPISTYNDOPYOYOHLIANTGGSKSTKYPAAYAGGYYYULHADYIPYRRG
DSRGSLLSPRPISTYNDOPYOYOHLIANTGGSKSTKYNAAYOFYOYOYOYYTHADYIPYRRG
DSRGSLLSPRPISTYNDOPYOYOHLIANTGGSKSTKYNAAYOFYOYOYOYYTHADYIPYRRG
DSRGSLLSPRPISTYNDOPYOYOHLIANTGGSKSTKYNAAYOFYOYOYOYOYOYOY
TOALAHGYOPRINTGOPTOYAHLIANTGGSKSTKYNAAYOFYOYOYOYOYOY
TOALAHGYOPRINTGOPTOYAHLIANTGGSKSTKYNAAYOFYOYOYOYOYOY
TOALAHGYOPRINTGOTTTGAPTTTGAPTTTYTOTGATTTMPOODAYSROGREGTTGG
RIGTYLLOQPATTATATATATPGSYYYYHPNITEVATGELYYOY
TOALAHGYTGODSYIDCNTCYTQYOYOTSIDPTTTITTTTMPOODAYSROGREGTGG
RIGTYREVPREERPSOMPOSSYILCEYDAGGANYYPHNITEVATTTMPGLYYOY
TOALAHGYTGODSYIDCNTCYTYOYOUNDYLTATTYTYINGAAOAPPPSSMOONWKC
LVRLKPTLGGPPTLYRLGAVONDYNTTTTYINGAAOAPPPSSAASAYTTGGTAG
ALAAYCLTTGSYYYYGGENDOWYLTTTYTYINGAAOAPPPSSAASAYTTGGTAAGGS
GLGKYYUDLAGYGAGVAGAVARALEGFWARMINFISGIOYLAGISTLPGN
ALAAYCLTTGGSYYYUGGSYANAYONGGENEYTYOFTYATGATGGATAAGGS
GLGKYYUDLLAGYGAGVAGAVARALEGFWARALEGFWARMINFISGIOYLAGISTLPGN
ALAAYCLTTGGAGVARAYAGAAAYOTSTAYUCANAGGA
ILRRHYGFGGEAAYOMNNRI IAFASRCHVYSFTHYVPESDAAARYTOILSSLTTOOLLK
RLHOWN NEOCSTFACTOORGALVARKWAGENEYTGGATAAGGSI
GLGKYYUDLLAGYGAGVAGALVARKWAGENEYTGFRYNGTGGATAYGGATGAGAAGGA
ILRRHYGFGGEAAYOWNNRI LAFASRCHVYSFTHYVPESDAAARYTOILSSLTTOOLLK
RLHOWN NEOCSTFACTOORGALVARKONGENERYTGGATARAGARRA
TAPAKATCHARAYONGAAAGEYYEVTGKYGTAGAGARARARA
                                                                                                                                                    ALTER OF THE TOTAL TOTAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ARGSPPSLASSSATQLSAPSLKATCTTCHNAPDADLIBANLLMRQEMGGNITRVESEN
KVVILDSFEPLRAEEDGREVSVPAEILRRTRKPPLALPIWARPDYNPPLLESWKDPDY
VPPVVHGCPLPPTKAPPIPPRRKRTVVLTESTVSSALAELATKAFGSSGSSAVDSGT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ATAPPDQPLDDGGTGSDVESYSSMPPLEGEPGDPDLSDGSWSTVSEEGGEDVVCCSMS
YTWTGALITPCAAEESKLPINALSNSLLRHHNLYYATTSRSASQRQKKVTFDRLQVLD
DHYRDVLKEWKAKASTVKARLLSVEEACKLTPPRSARSKFGYGAKDVRSLSSKAVHI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    HSVWKDLLEDTETPIQTTIMAKNEVFCVQPEKGGRKPARLIVPPDLGVRVCEKMALYD
VSTLLOQAWGSSYGPQYSPGORVEHLVNAMSKRNBWGFATDTRCFESTYTERDIRV
BESIYQCCDLAPEARLIRGLTENLYVGGPHINSKGONGYRRCASGVLTTSGGNT
TCYLKASAACRAAKLRDCTMLVCGDDLVVICESAGTQEDAASLRVFTEAMTRYSAPPG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   DPPOPEYDLELITSCSSNVSVAHDANNKRYYYLTRDPTIPLSRAAWETARHTPVNSWL
GNIIMYAPTLWARMILMTHFFSILLAQEQLEKALDCQIYGAIYSIEPLDLPQIIQRLH
GLSAFSLHTYSPGEINRVASCLRKLGVPPLRVWRHRARSVRAKLLSQGGRAAICGKYL
FNWAVRTKLKLTPIPPAASRLDLSSWFVAGYSGGDIYHSVSRARPRWFLMCLLLLSVGV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               9344
11
0
0
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Length:
Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                2655 g 1958
                                                               /product-"polyprotein"
/protein_id-"BAB18805.1"
/db_xref="GI:11559451"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Gaps:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            4.49
49.00
100.008
100.008
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             GIYLLPNR"
1 2842 C
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Percent Similarity:
Best Local Similarity:
Query Match:
```

0285-44-2111 x3334 0285-44-1557.

Phone: Japan

1 GlySerValValIleValGlyArgIleValLeu 11

US-09-965-594-26 (1-11) x AB049092 (1-9344)

1889

BASE COUNT

NIBING

Alignment Scores: Pred. No.:

Score:

Fax:

40000

9

source

FEATURES

5'UTR CDS

```
Hepatitis C virus polyprotein precursor, gene, complete cds. AF176573
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      VGDLCGSVFPVSQLFTESPRRHQTVQDCNCSIYPGHISGHRMAMDMMMMSPTAALVVV
SQLLRIPQAIVDMVAGAHWGVLAGLAYYSMVGNMAKVMIVLLLFAGVDGTTHTTGGAA
ARATQGFTSFFSLGPSQKIQLINTNGSWHINRTALNCNDSLQTGFLAALFYTYRFNAS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       /translation-"MSTNPKPQRKTKRNTNRRPQDVKPPGGGQIVGGVTLLPRRGPRL
GVRATRKTSERSQPRGRRQPIPKARHPEGRTWAQPGYPWPLYGNEGLGWAGWLLSPRG
SRPSWGPTDPRRRSRNIGKVIDTLTCGFADLMGYIPLVGAPLGGVARALAHGVRVLED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          AQGLIRACMLVRKAAGGHYIQMALMKLAALSGTYVYDHLTPLQDWAHAGLRDLAAAVE
PVVFSDWETKVITWGADTAACGDIISGLPVSARRGREILLGPADGLEEQGWRLLAPIT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             AYSQQTRGLLGCIITSLTGRDKNQVEGEVQVVSTATQSFLATCVNGVC#TVYHGACTK
TLAGPRGPITQMYTNVDQDLVGWQAPPGARSLTPCTCGSSDLYLVTRHADVIPVRRRG
DSRGSLLSPRPISYLKGSSGGPLLCPSGHAVGIFRAAVCTRGVAKAVDFVPVESWETT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        GVNYATGNIPGCSFSIFLLALLSCLTIPASAYEVRNVSGVYHVTNDCSNSSIVYEAAD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     VIMHTPGCVPCVQDGNTSRCWVALTPTLAARNASVPVTAIRRHVDLLVGTAAFCSAMY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           GCPERMASCRPIDKFDGGWGPITYAEPDSSDGRPYCWHYAPRPCGIVPASGVCGPVYC
FTPSPVVVGTTDRFGVPITTWGENETDVLLLINITRPPLGNWFGCTWMNSTGFIKTCGG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  PPCNIGGAGNTTLTCPTDCFRKHPEATYTKCGSGPWLTPRCLVDYPYRLWHYPCAVNF
TIFKVRMYVGGVEHRLNAACNWTRGERCDLEDRDRSELSPLLLSTTEWQVLPCSFTTL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            PALSTGLIHLHONIVDVQYLYGIGSAVIPFAIKWEYVLLLFLLLADARVCACLWMMLL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 iaqaeralenivvinaasvagahgiisfivffcaawy ikgrivpgaayalygvwpiil
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   LLLALPPRAYAMDREMAASCGGAVFVGLVLLTLSPHYKVFLARL IWWSQYFTTRAEAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     <u>LOVWVPPLNIRGGRDAJILLMCAVHSELIFDITKFLLAJLGPLMMFQARITRVPYFVR</u>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Molecular cloning HCV Russian isolate 1b from the serum of patient
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         //note="cleaved into C, E1, E2, p7, NS2, NS3, NS4a, NS4b, NS5a, and NS5b proteins" /codon_start=1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              2 (bases 1 to 9600)
Mokhonov,V.V., Samokhvalov,E.I., Novikov,D.V., Shatalov,A.G. and
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ., Samokhvalov, E.I., Novikov, D.V., Shatalov, A.G. and
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Direct Submission
Submitted (09-AUG-1999) Molecular Genetics, Ivanovsky Virology
Institute (09-AUG-1999) Molecular Genetics, Ivanovsky Virology
Institute (09-AUG-1999) Molecular Genetics, Ivanovsky Virology
Location/Cualifiers
1. 9600
/organism="Hepatitis C virus"
/mol_type="genomic RNA"
/strain="274933RU"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       /db_xref="taxon:11103"
/country="Russia"
/note="lsolated from acute hepatitis patient serum"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               /product="polyprotein precursor"
/protein_id="AAD50312.1"
/db_xref="G1:5738247"
                                     Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                       Matches:
                                                                                                                                                                                                                        US-09-965-594-26 (1-11) x HPCRNA (1-9456)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                AF176573.1 GI:5738246
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          /isolate-"lb"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Hepacivirus.

1 (bases 1 to 9600)

Mokhonov, V.V., Samokh
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              with acute hepatitis
   49.00
100.008
100.008
100.008
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   .341
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Prilipov, A.G
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Unpublished
                                     Percent Similarity:
Best Local Similarity:
                                                                                                               Query Match:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     DEFINITION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       5'UTR
CDS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ORGANISM
                                                                                                                                                                                                                                                                                                                                                                                                                                     RESULT 10
AF176573
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          TITLE
JOURNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    VERSION
KEYWORDS
SOURCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       AUTHORS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    REFERENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          AUTHORS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            JOURNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     REFERENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  FEATURES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         TITLE
                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                 ď
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              I AGAEAALENILVYLNARGSVAGVING ILSELVERGARVI IKERLVEGARVEGARVETURARGLILIALPPRAYAMAREMAASCGGAVFVGLALITISPYTKVETARLIWALOTITRAEALLILILALPPRAYAMAREMAASCGGAVFVGLALITISPYTKVETARLIWALOTITRAEAH LOVWI PPLAVYRGROBAI ILLICAAHPELIF POITYKLLIAI IEGPIMULOAGITKMPYEVRAGGITRAMUNKVAGGROTITGLAPARATIOTITRATAI IEGPIMULOAGITKMPYEVRAGGITRACMUNKVATVARGANTITAGYOTRALIGGITISLIGGENDANGVEGENOVNSTATOSFLATOVNOVENTYHGAGSK TLAGPKGPITOMYTNVOGDLVGWAPASGARSITPCTCGSSDIYLUTRYADVIPVRRAG DSRGSLLGGITISLIGGSHLAPSGARSITPCTCGSSDIYLUTRYADVIPVRRAG DSRGSLLGGOTTONYTNVOGDLVGWAPASGARSITPCTCGSSDIYLUTRYADVIPVRRAG DSRGSLLGGOTTONYTNVOGLVAHARTGSGKSTKVPAASTANGTKGVARANDFVPRESMETT MSSPVFTDNSSPRYDOTFOVAHLHAPTGSGKSTKVPAASAAGYKYLVANBSVAATLGGGSGGAYDII ICDECHS TDSTTIKGIGDPUTRTGVRTITTGAFTTYSTGKFLADGGCSGGAYDII ICDECHS TDSTTIKGIGDPUTRTGVRTITTGAFTTYSTGKFLADGGCSGGAYDII ICDECHS TDSTTIKGIGTUDOWSTRTGVRTITTGAFTTSTGKFLADGGCSGGAYDII ICDECHS TDSTTIKGIGTDVAYATRAEGSVYVPRANTESVALSTTGETPFYGK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      DHLEFWESVFTGLTHIDAHFLSGTKQAGDNFPYLVAYOATVCARAQAPPPSWDOWKC
LIRLKPTLHGPTPLLYRLGAVQNEVTLTHPITKYIMACMSADLEVYTSTWVLVGGYLA
ALAAYCLTGSVYIVGRIVLSGRPAYIPDREVLYQEFDEMESCGSHLPYIEOGMQLAE
PALAAYCLTGLOIATKQAFAYEVYENKARLETFWAKHWMFISGYQYLAGLSTLPGN
PALASLWAFTAKYYPLTTQSYLLFNILGGWYAAQLAPPSAAGAFVGAOIAGAAVGSI
GLGKVLVDILAGYGAGVAGALVAFKIMSGETPSAEDLVNLLPAILSPGALVVGVVCAA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   IORRHYGPGEGAVOMNNELIAFASRGNHYSPTHYVPESDAAARYTQILSSLTITOLLK
RLHOWINEDCSTPCSGSWLRDIWDWICSVLTDFKTWLOSKLLPRLPGVPFFSCQRGYK
GVWRGDGIMQTTCPCGAQITGHVKNGSMRIVGPKTCSNTWHGTFPINAYTTGPCTPSP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                YAPACKPVLREEVDEGVGLNQYPVGSQLPCEPBEDDVAVLTSMLTDPSHITAEAAKRRL
ARGSPPSLASSSASQLSAPSLKATCTTHHDSPDADLIBANLLMRQEMGGNITRVESEN
KVVILDSFEPLRAEEDEREVSVAAEIIRKTRRFPPAMPVWARPDYNPPLLESWKDPDY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     VPPVVHGCPLPPTKAPPIPPRKRTVVLTESTVSSALAELATKTFGSSGSSAVDSGT
ATAPPDQTSNDGDRESDAESYSSMPPLEGEPGDPDLSDGSWSTVSEEASGDVVCCSMS
YTWTGALITPCAAEESKLPINPLSNSLLRHHNMYYATTSRSASLRQKKVTFDRLQVLD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  GNIIMYAPTLWARMILMTHFFSILLAQEQLEKALECQIYGACYSIEPLDLPQIIERLH
GLSAFSLHSYSPGEINRVASCLRKLGVPPLRVWRHRARSVRAKLLSQGGRAATCGKYL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SRPSWGPNDPRRRSRNLGKVIDTLTCGFADLMGYIPLVGAPLGGAARALAHGVRALED
GVNYATGNLPGCSFSIFLLALLSCLTIPASAYEVRNRSGVYHVTNDCSNSSIVYEAAD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                MIMHNGGCVPCVRENNSSRCWVALTPTLAARNASVPTTTIRRHVDLLVGAAALCSAMY
VGDLCGSVFLVSQLFTFSPRRHETVQDCNCSIYPGHITGHRMAMDMMMMSPTTALVV
SQLLRIPQAVMDMVAGAHWGVLAGLAYYSMVGNWAKVLIVLLLFAGVDGNTYVTGGAA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ARGASGITSLFSRGPSQKIQLVNTNGSWHINRTALNCNDSFNTGFLAALFYAHRFNSS
GCPERMASCRSIDKFDQGWGPITYYQGDSPDQRPYCWHYPPRRCGIVPASEVCGPVYC
FTPSPVVVGTTDRLGVPTYNWGENETDVLLLNNTRPPQGNWFGCTWMNTTGFTKTCGG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             PPCNIGGAGNNTLICPTDCFRKHPEAIYTKCGSGPWLIPRCLVDYPYRLMHYPCTVNF
TIFKVRMYVGGVEHRLDAACNWTRGERCALEDRDRSELSPLLLSITEWQILPCSFTTL
PALSTGLIHLHRNIVDVQYLYGIGSAVVSFAIKWEYVLLLFLLLADARVCACLMMALL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               TDALMTGFTGDFDSVIDCNTCVTQTVDFSLDPTFTIETTTVPQDAVSRSQRRGKTGRG
RRGIYRFVTPGERPSGMFDSSVLCECYDAGCAMYELTPAETSVRLRAYLNTPGLPVCQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          rsvwkdlledtetpidttimaknevfcvqpekggrkparlivfpdlgvrvcekmalyd
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ICYLKASAACRAAKLQDCTMLVNGDDLVVICESAGTQEDAASLRVFTEAMTRYSAPPG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                          /translation-"mstnpkpqrktkrntnrrpqdvkfpgggjvggvyllpprrgpr
gvratrktsersqprgrrqpipkarrpegramaqpgypwplygnegmgwagullsprg
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               APNYSRALWRVAAEEYVEVTRVGDFHYITGMTTONIKCPCOVPAPEFFTEVDGVRLHR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    DPPQPEYDLELITSCSSNVSVAHDASGKRVYYLTRDPTTPLARAAWETARHTPVNSWL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            FNWAVRTKLKLTPIPAASRLDLSGWFVAGYGGGDIYHSLSRARPRWFMLCLLLLSVGV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 DHYRDVLKEMKAKASTVKAKLLSIEEACKLTPPHSAKSKFGYGAKDVRNLSSKAVNHI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             eesiyoccdlapeakoaikslierlyiggplinskgoncgyrcrasgvlitscgntl
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        VVSTLPQAVMGSSYGFQYSPGQRVEFLVNTWKSKKCPMGFSYDTRCFDSTVTENDIR
                                                                                                                                                                                                                                                  342. .9374
/note-"C, E, NS1/E2, NS3, NS4 and NS5*
/codon_start-1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    9456
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 a T-stretch of 45 bp."
2668 g 2068 t
                                                                                                                                                                                                                                                                                                                                                              /product-"polyprotein precursor"
/protein_id-"BAA01728.1"
/db_xref-"GI:471117"
                                                                      /organism="Hepatitis C virus"
/mol_type="genomic RNA"
/isolate="HC-C2"
                                                                                                                                                                                      'db_xref-"taxon:11103"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      2668 g
Location/Qualifiers
1. .9456
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      GIYLLPNR"
9375. 9456
/note="with a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      2811 c
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       4.55
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1909 a
```

Alignment Scores:

BASE COUNT

ORIGIN

Pred. No.:

```
Percent Similarity:
                                                   Query Match:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                5'UIR
CDS
                                                                                                                                                                                                                                                                                                                                                           DEFINITION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ORGANISM
                                                                                                                                                                                                                                                                                                                                                                                     ACCESSION
VERSION
KEYWORDS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  REFERENCE
AUTHORS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   MEDLINE
PUBMED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              AUTHORS
TITLE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 JOURNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           JOURNAL
                                                                                                                                                                                                                                                                                     RESULT 12
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         REFERENCE
                                                                                                                                                                                                                                                                                                          AF268278
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    FEATURES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          TITLE
                                                                                                                                                                                                                               g
                                                                                                                                           THE PROPERTY PROPERTY OF THE TOTAL OF THE STATE OF THE ST
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        PAT 20-APR-2002
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   FNWAVRTKLKLTPIPAASQLDLSGWFVAGYGGGDIYHSLSRARPRWPMLCLLLLFVGV
GIYLLPNR"
9375. ,9600
                                                                        AIPIEIIKGGRHLIFCHSKKRCDELAAKLSGLGLNAVAYYRGLDVSVIPTSGNVVVVA
TDALMTGFTGDFDSVIDCNTCVTQTVDFSLDPTFTIETTTVPQDAVSRSQRRGRTGRG
RRGIYRFVTPGERPSGMFDSSVLCECYDAGCAMYELTPAETSVRLRAYLNTPGLPVCQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1 (bases 1 to 12734)
Hong,Z., Lai,W.C.H. and Lau,J.Y.N.
Hong,Z., Lai,W.C.H. and Lau,J.Y.N.
Hepatitis C virus protease-dependent chimeric pestivirus
Patent: US 6326137-A 1 04-DEC-2001;
Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        linear
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       12734
11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     9600
11
0
0
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Length:
Matches:
Conservative:
Mismatches:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1 GlySerValVallleValGlyArglleValLeu 11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Length:
Matches:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      '5. .9600
2887 c 2696 g 2100
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Sequence 1 from patent US 6326137.
AR179057
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Gaps:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    US-09-965-594-26 (1-11) x AF176573 (1-9600)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              /organism="unknown"
2604 c 3295 q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 AR179057.1 GI:20220612
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     4.62
49.00
100.00%
100.00%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       6.27
49.00
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Unclassified
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Score:
Percent Similarity:
Best Local Similarity:
Query Match:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 4032 a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Unknown.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Unknown.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1917
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                source
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Pred. No.:
Score:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 BASE COUNT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   DEFINITION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ORGANISM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  TITLE
JOURNAL
FEATURES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Pred. No.:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ACCESSION
VERSION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 BASE COUNT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        RESULT 11
AR179057
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          AUTHORS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  REFERENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         KEYWORDS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ORIGIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     a
```

```
LOCHEDATE UNIVERSITY IN PRINCED WEIGHT LEGIN IN THE CHARLES ALL IN LOCK TO REPORT WHEN THE PRINCES ALL IN LOCK TO REPORT CHARLES AND THE CHARL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Pestivirus type 1
Postivirus type 1
Viruses; SSRNA positive-strand viruses, no DNA stage; Flaviviridae;
Viruses; SSRNA positive-strand viruses, no DNA stage; Flaviviridae;
Viruses; SSRNA positive-strand
1 (bases 1 to 12734)
1ai,v.C., Dong,W., Skelton,A., Ingravallo,P., Vassilev,V.,
Generation and characterization of a hepatitis C virus NS3
protease-dependent bovine viral diarrhea virus
1. virol. 74 (14), 6339-6347 (2000)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               2 (bases 1 to 12734)
Lai, V.C.H. and Hong, Z.
Direct Submission
Submitted (16-MAY-2000) Antiviral Therapy, Schering-Plough Research Institute, 2015 Galloping Hill Road, Kenilworth, NJ 07033-0539, USA Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SYSDTKEEGATKKRTOKPDKLERGKMKIVPKESEKDSKTKPPDATIVVEGVKYOVRKK
GKTKSKNTQDGLYHNKNKPQESRKKLEKALLAMAIIAIVLFQYTMGENITQMNLQDNG
TEGIQRAMFQRGVNRSLHGIMPEKICTGVPSHLATDIELKTIHGMADASEKTNYTCCR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           BGPVEKTACT FNY TKTLKNKY FEPRDSY FQQYMLKGEY QYWFDLEVTDHHRDY FAES I
LVVVVALLGGRYVLWLLYTYNVLSEQKALGIQY GSGEVVMGNLLTHNNI EVVTY FLL.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   /translation="MELNTNEGSGSVVIVGRIVLSGSGSITACAOOTRGLLGCKITSL
TARRNAW DEGENGIVSTATOPTELATOL NOVCWPTY VEAGATTA LASPROVIONTIND
QDIVGWPAPOGSRSITPCTGSSDLYLVTRHANVIPVRRRGDSRGSLLSPRPISVIKG
SSGGPLLCPAGHAVGLFRAAVCTRGVAKAVDFIPVENLETTTRSGSGADTEDVVCCSM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          TVVLIVIĞLIIARRDPTIVPLVTIMAALRVTELTHQPGVDIAVAVMTITLIMVSYVTD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      LTLDFMYYMHRKVIEEISGGTNIISRLVÄALIELMASMEEEESKGLKRFYLLSGRLRN
LIIKHKVRNETVASWYGEEEVYGMPKIMTIIKASTLSKSRHCIICTVCEGREWKGGTC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               VRL 12-JUL-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         LYLLEREESVKKWYLLLYHILVVHPIKSVIVILLMIGDVVKADSGGOEYLGKIDLCFT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 YFRYKKWLOCILSLVSGVFLIRSLIYLGRIEMPEVTIPNWRPLTLILLYLISTTIVTR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               IDTTRVDSIYDVDESGEGVYLFPSRQKAQGNFSILLPLIKATLISCVSSKWQLIYMSY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     linear
           0000
           Conservative:
                                                                                                                                                                                                                                                                                                                                                                                                                                  /organism="Pestivirus type 1"
/mol_type="genomic RNA"
/db_xref="taxon:11099"
1. 385
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               RNA
                                                            Mismatches:
                                                                                                                                                                                                                                                                                                                                                                           1 GlySerValValIleValGlyArgIleValLeu 11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Pestivirus type 1, complete genome. AF268278
                                                                                                                   Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               /product-"polyprotein"
/protein_id-"AAF82566.1"
/db_xref-"G1:9049957"
                                                                                                                                                                                                                                                                        US-09-965-594-26 (1-11) x AR179057 (1-12734)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          12734 bp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             /codon_start=1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    AF268278.1 GI:9049956
100.00%
100.00%
100.00%
                                                       Best Local Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        10864644
```

```
Unclassified.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Unclassified.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Percent Similarity:
Best Local Similarity:
Query Match:
DB:
                                                                                                                                                                                                                      Percent Similarity:
Best Local Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                         AR145221
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Unknown.
                                                                                                              91
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Unknown
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Unknown
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Unknown
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Alignment Scores:
                                                                                                                                                                   Alignment Scores:
                                                                                                                                                                                                                                                       Query Match:
DB:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  source
                                                                                                                                                                                                                                                                                                                                                                                                                                                                        DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
                                                                                                          BASE COUNT
ORIGIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                BASE COUNT
ORIGIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                DEFINITION
                                                                                                                                                                                Pred. No.:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ORGANISM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Pred. No.:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ORGANISM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   REFERENCE
AUTHORS
TITLE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           RESULT 15
AR145197
LOCUS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ACCESSION
VERSION
KEYWORDS
                                                                                                                                                                                                                                                                                                                                                                                                                    RESULT 14
AR14521
LOCUS
                                  JOURNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            JOURNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 REFERENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  AUTHORS
                                                       FEATURES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               FEATURES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 TITLE
TITLE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SOURCE
                                                                                                                                                                                                       Score:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                 HYYKCVTAKIDYSKGKMLLATDKWEVEHGVITRIAKRYTGVGFNGAYLCDEPNIRALV
ERDCATITKNTVQFLKMKKGCAFTYDLTISNLTRLIELVHRNNLEEKEIPTATVTTWL
ATTVNEDVOTIRPVLGERVIPDPVYDINLQPENQYDTSEVGITIIGREFLMTTGVTP
VLEKVEPDASDNGNSVKIGLDEGNYPORIQTHINRDARPFIMILGSRNSIS
NRAKTARININIYTGNPREIRDLMAGRALVVALRDVOPELSENVDFKGTFLDREALE
ALSLGQPKPKQVTKEAVRNIEDGKKDVEIPNWPASODPVFLEVALKNDKYYLVGDVGE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          VKDOAKALGATDOTRIIKEVGSRTYAMKLSSWFLQASNKOMSLTPLFEELLLRCPPAT
SKNCHMASYOLAGOWBEPLGGGWHGTIPARKYHIPTSATLKLKDFIEEERKRPR
VKOTWIREHWKWILKIRFOGMLWIRKMLNPCKLSEQLDREGRRRIYWHOIGTIMS
AGIRLEKLPIVRAQTDTKTFHEAIRDKIDKSENRONPELHNKLLEIFHTIAQPTLKHT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    PAT 08-AUG-2001
                                                                    TGWAYTHQGGISSVDHVTAGKDLLVCDSMGRTRVVCQSNNRLTDETEYGVKTDSGCPD
GRACYVLNPRANISGSKGAVVHLQRYGGEFTCVTASGTPERTDLKNIKGWSGLDIFE
ASSGRVVGRVKVGRYKCOGITGVSKNTADLTEWVKKITSWNRGDFKOITLA
TGAGKTTELPRAVIEEIGRHKRVLVLIPLRAAAESVYQYMRLKHPSISFNLRIGDWKL
                                                                                                                                            GDMATGITYASYGYFCQMPOPKLRAAWVEYSYIFLDEYHCATPEQLAIIGKIHRFSES
IRVVAMTATPAGSVTTTGQKHPIEEFIAPEVMKGEDLGSQFLDIAGLKIPVDEMKGNM
LVFVPTRNMAVEVAKKLKAKGYNSGYYSGEDPANLRVVTSQSPYVIVATNAIESGVT
                                                                                                                                                                                                LPDLDTVIDTGLKCEKRVRVSSKIPFIVTGLKRMAVTVGEQÄQRRGRVGRVKPORYYR
SQETATGSKDYHYDLLQAQRYGIEDGINVTKSFREMNYDWSLYEEDSLLITQLEILINN
LLISEDLPAAVKNIMARTDHPEPIQLAYNSYEVQVPVLFPKIRNGEVTDTYENYSFLN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          IEFCSHTPVPVRWSDNTSSHMAGRDTAVILSKWATRLDSSGERGTTAYEKAVAFSFLL.
MYSWNPLVRRICLLVLSQQPETDPSKHATYYYKGDPIGAYKDVIGRNLSELKRTGFEK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    PKNEKRDVSDDWQAGDLVVEKRPRVIQYPEAKTRLAITKVMYNWKQQPVVIPGYEGK
TPLFNIFDKVRKEWDSFNEPVAVSFDTKAWDTQVTSKDLQLIGEIQKYYYKKEWHKFI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SFNRVARIHVCGDDGFLITEKGLGLKFANKGMQILHEAGKPQKITEGEKMKVAYRFED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                LANLNLSLSTLGIWTKHTSKRIIQDCVAIGKEEGNWLVNADRLISSKTGHLYIPDKGF
                                                     GWILRGPAVCKKITEHEKCHINILDKLTAFFGIMPRGTTPRAPVRFPTSLLKVRRGLE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      YGEVTWEQLEAGINRKGAAGFLEKKNIGEVLDSEKHLVEQLVRDLKAGRKIKYYETAI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          DT ITDHMTEVPVITADGEVY I RNGQRGSGQPDTSAGNSMLNVLTMMYAFCESTGVPYK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 TLOGKHYEQLQLRTETNPVMGVGTERYKLGPIVNLLLRRLKILLMTAVGVSS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Unknown.
Unclassified.
1 (bases 1 to 75)
Malcolm,B.A., Taremi,S.Shane., Weber,P.C. and Yao,N.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1 others
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    linear
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        12734
11
0
0
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Length:
Matches:
Conservative:
Mismatches:
Indels:
Gaps:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1 GlySerValValIleValGlyArgIleValLeu 11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Sequence 48 from patent US 6211338.
AR145217
AR145217.1 GI:15107084
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (1-12734)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  75
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   3293 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     US-09-965-594-26 (1-11) x AF268278
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        6.27
49.00
100.00%
100.00%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                12509. .127
1 2608 c
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Percent Similarity:
Best Local Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                AR145217
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Unknown.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     4030
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Query Match:
DB:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   DEFINITION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ORGANISM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  3'UTR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Pred. No.:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            RESULT 13
AR145217
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   BASE COUNT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ACCESSION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   REFERENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      AUTHORS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        KEYWORDS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         VERSION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ORIGIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SOURCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
```

```
PAT 08-AUG-2001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                08-AUG-2001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Malcolm, B.A., Taremi, S.Shane., Weber, P.C. and Yao, N. Single-chain recombinant complexes of hepatitis C virus NS3 protease and MAA cofactor peptide Patent: US 621338-A 55 03-ARR-2001;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1 (bass: 1 to 78)
Malcolm,B.A., Taremi,S.Shane, Weber,P.C. and Yao,N.
Single-chain recombinant complexes of hepatitis C virus NS3
protease and N84A cofactor peptide
Patent: US 6211338-A 26 03-APR-2001;
Single-chain recombinant complexes of hepatitis C virus NS3 protease and NS4A cofactor peptide Patent: US 6211338-A 48 03-APR-2001;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                PAT
                                                                                                                                                                                                                                                                                                                                                            linear
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                linear
                                                                                                                                               10
10
0
0
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  001000
                                                                                                                                            Length:
Matches:
Conservative:
Mismatches:
Indels:
Gaps:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Matches:
Conservative:
Mismatches:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                DNA
                                                                                                                                                                                                                                                                                      DNA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1 GlySerValValileValGlyArgileValLeu 11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1 GlySerValValIleValGlyArglleValLeu 11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ب
                                                                                       نډ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Sequence 26 from patent US 6211338. ARI45197
                                                                                                                                                                                                                                                                                                                                                                       Sequence 55 from patent DS 621[338.
AR145221
AR145221.1 GI:15107088
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         33
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Length:
                                                                                       30
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Gaps:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                US-09-965-594-26 (1-11) x AR145221 (1-75)
                                                                                                                                                                                                                                                (1-75)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      /organism="unknown"
12 c 16 q
                                           Location/Qualifiers
1...75
                                                                       /organism="unknown"
14 c 15 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Location/Qualifiers
                                                                                       б
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        AR145197.1 GI:15107064
                                                                                                                                                                                                                                             US-09-965-594-26 (1-11) x AR145217
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            48.00
100.00%
90.91%
97.96%
                                                                                                                                                           48.00
100.00%
90.91%
97.96%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       JOURNAL
```

```
78
10
1
0
0
                                                        Length:
Matches:
Conservative:
Mismatches:
Indels:
Gaps:
                          31 t
Location/Qualifiers
1. .78
/organism-"unknown"
2 11 c 18 9
                                                       0.0426
48.00
100.00%
90.91%
                                                Alignment Scores:

Pred. No.:

Score:

Percent Similarity: 10

Best Local Similarity: 90

Query Match: 67
                        BASE COUNT 18 3
FEATURES
SOUICE
```

Search completed: August 31, 2003, 00:46:50 Job time : 151.976 secs

US-09-965-594-26 (1-11) x AR145197 (1-78)

δ q

```
Hepatitis C virus
Hepatitis C virus
HCV NS4A DNA fragm
HCV peptide C14-1.
HCV peptide C14-1.
HCV Deptide C14-1.
HCV I chimeric ep
Hepatitis C virus
HCV I chimeric ep
Hepatitis C virus
HCV I chimeric ep
Non-A, Non-B Hepat
Non-A, Non-B Hepat
HCV antigen. Synt
DNA encoding Hepat
HCV antigen. Synt
DNA associated wit
Anti-viral synthet
HCV NS4A-NS3 compl
                                                                                                                                                                    Chimeric BVDV/HCV
HCV NS4A-NS3 compl
                                                                                                                            Hepatitis C v
Hepatitis C v
HCV NS2-NS4 c
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              protease; viral replication; chronic liver disease; liver cancer; primer; ss.
                                                                        Description
and is derived by analysis of the total score distribution
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 HCV NS4A-NS3 fusion protease oligonucleotide #1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ALIGNMENTS
                             SUMMARIES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   AA250043
AAQ27012
                                                                                                                                                                                                                                                                                                                                                                                                                                   AAQ25753
AAQ62690
                                                                                               AAA73336
                                                                                                                                                                                                                                                                                                                                                                                            AAQ23456
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  AAX80363
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               4AX80348
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          AAX80350
                                                                                                                                                                       ABA95615
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ABX157
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                BP
                                                                                                 AAA73336 standard; DNA; 68
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       (first entry)
                                                                     Length
                                        %
Query
Match I
                                                                                                              Hepatitis; NS3
liver failure;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   40200040707-A1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      19-DEC-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Synthetic
                                                                                                                                                                                                                                                                                 Score
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          AAA73336;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   RESULI 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   AAA73336
                                                                                                                                                                                                                                                                                              Result
                                                                       Š.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              August 30, 2003, 19:13:57; Search time 10.2149 Seconds (without alignments) 2906.924 Million cell updates/sec
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         N. Geneseq_19Jun03:*

| SIDSI/gcgdata/geneseqn-embl/NA1980.DAT:*
| SIDSI/gcgdata/geneseqq-embl/NA1980.DAT:*
| SIDSI/gcgdata/geneseqq-embl/NA1981.DAT:*
| SIDSI/gcgdata/geneseqq-embl/NA1981.DAT:*
| SIDSI/gcgdata/geneseqq-embl/NA1984.DAT:*
| SIDSI/gcgdata/geneseqq-embl/NA1986.DAT:*
| SIDSI/gcgdata/geneseqq-embl/NA1986.DAT:*
| SIDSI/gcgdata/geneseqq-embl/NA1986.DAT:*
| SIDSI/gcgdata/geneseqq-embl/NA1986.DAT:*
| SIDSI/gcgdata/geneseqq-embl/NA1989.DAT:*
| SIDSI/gcgdata/geneseqq-embl/NA1989.DAT:*
| SIDSI/gcgdata/geneseqq-embl/NA1989.DAT:*
| SIDSI/gcgdata/geneseqq-embl/NA198.DAT:*
| SIDSI/gcgdata/geneseqq-embl/NA1999.DAT:*
| SIDSI/gcgdata/geneseqq-embl/NA2000.DAT:*
| SIDSI/gcgdata/geneseqq-embl/NA2001.DAT:*
| SI
                                                                                                                                                                                                                                                                                                                      5105512
           GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.
                                                                  nucleic search, using frame_plus_p2n model
                                                                                                                                                                                                                                                                                                                     of hits satisfying chosen parameters:
                                                                                                                                                                                                                                                                                         2552756 segs, 1349719017 residues
                                                                                                                                                                                                                                                                                                                                                                                     Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries
                                                                                                                                                                                                                        0.5
0.5
7.0
                                                                                                                                                                                                                                      , Ygapext
, Fgapext
, Delext
                                                                                                                                                                                                                                                                                                                                               length: 0
length: 2000000000
                                                                                                                                                  US-09-965-594-26
49
                                                                                                                                                                               1 GSVVIVGRIVL 11
                                                                                                                                                                                                        BLOSUM62
Xgapop 10.0 , X
Ygapop 10.0 , Y
Fgapop 6.0 , F
Delop 6.0 , F
                                                                                                                                                                                                                                                                                                                                                 sed
                                                                                                                                                                                                           Scoring table:
                                                                                                                                                                  Perfect score:
                                                                                                                                                                                                                                                                                                                      Total number
                                                                                                                                                                                                                                                                                                                                               Minimum DB s
Maximum DB s
                                                                  OM protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Database :
                                                                                                                                                                                 Sequence:
                                                                                                                                                                                                                                                                                            Searched:
                                                                                              Run on:
```

Goldfarb

Zhang Y,

~

```
Wittekind M, Weinheimer S,
                                                                                                                                                                                                                                                                                                                                                             Query Match:
                                                                                                                                                                                                                                                                                                                                                                                                                                                               RESULT 3
   õ
                                                                                                                                                                                                                                                                                                                                                                                                                                         윱
                                                                                                                                                                                                                             linker segments which were used in the creation of a fusion molecule of the coding sequences for Hepatitis C virus (HCV) NS3 and NS4A protease enzymes. These proteins are poth essential for the replication of the virus, acting to cleave its replicative proteins from the polyprotein produced from the HCV genome. Inhibitors of the two proteins should be effective as antiviral treatments of HCV infection. This is useful as HCV can lead to chronic liver disease such as cirrhosis, liver failure and NS3-NS4A fusion proteins which can be used to identify inhibitors of this type, as well as enabling structural studies of the protease and
                                                                                                                                     Modified hepatitis C virus (HCV) NS3 protease comprising at least 1 substitution of a hydrophobic alpha-helix 0 amino acid to a hydrophilic amino acid, useful for screening inhibitors that may treat hepatitis C
                                                                                                                                                                                                                    present sequence is one of two oligonuclectides coding for NS4A and
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      protease; viral replication; chronic liver disease;
                                                                                                                                                                                                                                                                                                                                                                                                      00000
                                                                                          Goldfarb V;
                                                                                                                                                                                                                                                                                                                                                                                                                 Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                      Sequence 68 BP; 20 A; 14 C; 17 G; 17 T; 0 other;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                HCV NS4A-NS3 fusion protease oligonucleotide #2.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Length:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   liver failure; liver cancer; primer; ss
                                                                                          Weinheimer S, Zhang Y,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    US-09-965-594-26 (1-11) x AAA73336 (1-68)
                                                                                                                                                                                            Example 2; Page 22; 66pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          8
                                                                      (BRIM ) BRISTOL-MYERS SQUIBB CO
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (BRIM ) BRISTOL-MYERS SQUIBB
                                                                                                                                                                                                                                                                                                                                                protease; inhibitor complexes
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            06-JAN-2000; 2000WO-US00345
                         2000WO-US00345
                                               99US-0115271
                                                                                                                                                                                                                                                                                                                                                                                                    0.106
49.00
100.008
100.008
100.008
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            AAA73337 standard; DNA; 71
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          (first entry)
                                                                                                                 WPI; 2000-465976/40.
                                                                                                                                                                                                                                                                                                                                                                                                                            Percent Similarity:
Best Local Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Hepatitis; NS3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              WO200040707-A1
                                                                                          Wittekind M,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   08-JAN-1999;
                        06-JAN-2000;
                                                                                                                                                                                                                                                                                                                                                                                           Alignment Scores:
                                               08-JAN-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         19-DEC-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     13-JUL-2000
 13-JUL-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Synthetic.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   AAA73337;
                                                                                                                                                                                                                                                                                                                                                                                                                                                   Query Match:
DB:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 RESULT 2
                                                                                                                                                                                                                                                                                                                                                                                                                  Score:
a
```

```
The present sequence is one of two oligonuclectides coding for NS4A and linker segments which were used in the creation of a fusion molecule of the coding sequences for Hepatitis C virus (HCV) NS3 and NS4A protease enzymes. These proteins are both essential for the replication of the virus, acting to cleave its replicative proteins from the polyprotein produced from the HCV genome. Inhibitors of the two proteins should be effective as antiviral treatments of HCV infection. This is useful as HCV can lead to chronic liver disease such as cirrhosis, liver failure and liver cancer. The present invention concerns a number of NS3 mutants and NS3-NS4A fusion proteins which can be used to identify inhibitors of this
                                      Modified hepatitis C virus (HCV) NS3 protease comprising at least 1 substitution of a hydrophobic alpha-helix 0 amino acid to a hydrophilic amino acid, useful for screening inhibitors that may treat hepatitis C
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Hepatitis; NS3 protease; viral replication; chronic liver disease;
liver failure; liver cancer; ds.
                                                                                                                                                                                                                                                                                                                                                             type, as well as enabling structural studies of the protease and protease inhibitor complexes.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Hepatitis C virus NS4A-NS3 fusion protease coding sequence #1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               /*tag- a
/product- "NS3-NS4A fusion protein"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  0000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Goldfarb V;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Length:
Matches:
Conservative:
Mismatches:
                                                                                                                                                                                                                                                                                                                                                                                                                      Sequence 71 BP; 17 A; 18 C; 15 G; 21 T; 0 other;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1 GlySerValValIleValGlyArgIleValLeu 11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  2hang Y,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (1-71)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Location/Qualifiers
                                                                                                                                  Example 2; Page 23; 66pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          (BRIM ) BRISTOL-MYERS SQUIBB CO
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      BP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 US-09-965-594-26 (1-11) x AAA73337
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Weinheimer S,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      AAA73328 standard; DNA; 588
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      990S-0115271
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  06-JAN-2000; 2000WO-US00345
                                                                                                                                                                                                                                                                                                                                                                                                                                                                              0.111
49.00
100.00%
100.00%
100.00%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            19-DEC-2000 (first entry)
WPI; 2000-465976/40.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Percent Similarity:
Best Local Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Hepatitis C virus.
Synthetic.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     WO200040707-A1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Wittekind M,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      08-JAN-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                            Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            13-JUL-2000.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          AAA73328;
```

m

Page

```
8
                                                                                              The present sequence is the coding sequence for a fusion protein created using the Hepatitis C virus (HCV) NS3 and NS4A protease enzymes. These proteins are both essential for the replication of the virus, acting to cleave its replicative proteins from the polyprotein produced from the HCV genome. Inhibitors of the two proteins should be effective as antiviral treatments of HCV unfection. This is useful as HCV can lead the present invention concerns a number of NS3 mutants and NS3-NS4A fusion proteins which can be seed to identify inhibitors of this type, as well as enabling structural studies of the protease and
                            Modified hepatitis C virus (HCV) NS3 protease comprising at least 1 substitution of a hydrophobic alpha-helix 0 amino acid to a hydrophilic amino acid, useful for screening inhibitors that may treat hepatitis C
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Hepatitis; NS3 protease; viral replication; chronic liver disease;
liver failure; liver cancer; mutant; muteln; ds.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Hepatitis C virus NS4A-NS3 fusion protease coding sequence #2.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   /*tag= a
/product= "NS4A-NS3 fusion protein #2"
                                                                                                                                                                                                                                                    0
0
0
0
0
0
                                                                                                                                                                                                                      Sequence 588 BP; 97 A; 183 C; 153 G; 155 T; 0 other;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Goldfarb V;
                                                                                                                                                                                                                                                    Length:
Matches:
Conservative:
                                                                                                                                                                                                                                                                                                                                                      13 GGTTCCGTTGTTATCGTCGCCCGTATAGTACTG 45
                                                                                                                                                                                                                                                                                                                                            1 GlySerValValIleValGlyArglleValLeu 11
                                                                                                                                                                                                                                                                                 Mismatches:
                                                                                                                                                                                                                                                                                           Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Zhang Y,
                                                                                                                                                                                                                                                                                                                         US-09-965-594-26 (1-11) x AAA73328 (1-588)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Location/Qualifiers
                                                                              Disclosure; Fig 10; 66pp; English
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                (BRIM ) BRISTOL-MYERS SQUIBB CO.
                                                                                                                                                                                                                                                                                                                                                                                                       AAA73329 standard; DNA; 588 BP
                                                                                                                                                                                                    protease:inhibitor complexes.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Wittekind M, Weinheimer S,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        06-JAN-2000; 2000WO-US00345
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             99US-0115271
                                                                                                                                                                                                                                                    1.21
49.00
100.00%
100.00%
                                                                                                                                                                                                                                                                                                                                                                                                                                              19-DEC-2000 (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        WPI; 2000-465976/40.
P-PSDB; AAB15220.
2000-465976/40
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Hepatitis C virus.
                                                                                                                                                                                                                                                             Score:
Percent Similarity:
Best Local Similarity:
          P-PSDB; AAB15212
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 WO200040707-A1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             08-JAN-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     13-JUL-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Synthetic
                                                                                                                                                                                                                                                                                                                                                                                                                           AAA73329;
                                                                                                                                                                                                                                                                                           Query Match:
                                                                                                                                                                                                                                                                                                                                                                                              AAA7332
                                                                                                                                                                                                                                                                                                                                                                                    RESULT
                                                                                                                                                                                                                                                                                                                                                                                                        a
 8
```

```
fusion protein created using the Hepalitis C virus (HCV) NS3 and NS4A protease enzymes. These proteins are both essential for the replication of the virus, acting to cleave its replicative proteins from the polyprotein produced from the HCV genome. Inhibitors of the two proteins should be effective as antivirial treatments of HCV infection. This is useful as HCV can lead to chronic liver disease such as cirrhosis, liver failure and liver cancer. The present invention concerns a number of NS3 mutants and NS3-NS4A fusion proteins which can be used to identify inhibitors of this type, as well as enabling structural studies of the protease and protease:inhibitor complexes. The protein produced from this
Modified hepatitis C virus (HCV) NS3 protease comprising at least 1 substitution of a hydrophobic alpha-helix 0 amino acid to a hydrophilic amino acid, useful for screening inhibitors that may treat hepatitis {\sf C}
                                                                                                                                                                      present sequence is the coding sequence for a mutated version of a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Э
Н
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Clone; polypeptide; NS2-NS4; Hepatitis C; Virus; HCV; serum; transcriptase; cDNA; primer; allele; ss.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Takahashi K;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Sequence 588 BP; 103 A; 180 C; 156 G; 149 T; 0 other;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  588
111
0
0
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Conservative:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Mismatches:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            13 GGATCCGTTGTTATCGTCGCCGTATAGTACTG 45
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1 GlyServalvallleValGlyArgileValLeu 11
                                                                                                                                                                                                                                                                                                                                                                                                                               sequence contains the alpha-helix0-1 variant.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Length:
Matches:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Sek1 M,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Gaps:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        US-09-965-594-26 (1-11) x AAA73329 (1-588)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Honda Y, Murakami T,
                                                                                                                        Claim 26; Fig 12; 66pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ВР
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            (MITU ) MITSUBISHI KASEI CORP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     91JP-0172794.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 92EP-0109812
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            91JP-0139268
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     91JP-0332329
92JP-0099957
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1.21
49.00
100.00%
100.00%
100.00%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (updated)
(first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ø
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 HCV NS2-NS4 clone N13-1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              AAQ32477 standard; DNA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           9..608
/*tag=
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Percent Similarity:
Best Local Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           25-MAR-2003
26-APR-1993
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               11-JUN-1992;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     12-JUL-1991;
07-OCT-1991;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               11-JUN-1991;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     16-DEC-1991;
20-APR-1992;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  16-DEC-1992
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    EP518313-A2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Hayashi N,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             AAQ32477;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Query Match:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            AA032477
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     RESULT
```

Teranishi Y;

```
The sequences given in AA032472-82 and AA032442 are various clones which encode the NS2-NS4 regions of the Hepatitis C Virus (HCV) gene of the invention. These sequences were isolated from the serum of a patient suffering from hepatitis C (HC). The isolated RNA sequences were converted into cDNA using transcriptase in the presence of one of the primer sequences given in AA03253-164. The sequences were then amplified using primer pairs. The cDNA sequences isolated represent different alleles of the same region of the HCV gene. Sequence comparisons of these clones showed that it is possible for a AA032436.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Nucleic acid construct encoding chimeric Hepatitis C Virus (HCV) pestivirus genome where the Npro protease gene is replaced with NS3 protease gene, useful for in vivo screening of compounds which inhibit
                                                   New hepatitis C virus gene and its encoded protein - used for diagnosing and vaccinating against hepatitis C virus infections
                                                                                                                                                                                                                                                                                                            Sequence 618 BP; 135 A; 184 C; 176 G; 123 T; 0 other;
                                                                                                                                                                                                                                                                                                                                                        618
00
00
0
                                                                                                                                                                                                                                                                                                                                                                                 Conservative:
Mismatches:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  351 GGCAGCGTGGTCATTGTGGCAGGATCGTCTTG 383
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Pestivirus; Npro; protease; NS3; screening; ds
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1 GlySerValValIleValGlyArgIleValLeu 11
                                                                                                                                                                                                                                                                                   (Updated on 25-MAR-2003 to correct PN field.)
                                                                                                                                                                                                                                                                                                                                                        Length:
Matches:
                                                                                                                                                                                                                                                                                                                                                                                                              Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Columns 17-28; 20pp; English
                                                                                                Disclosure; Page 125-26; 305pp; English
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Chimeric - Bovine viral diarrhea virus.
Chimeric - Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                        US-09-965-594-26 (1-11) x AAQ32477 (1-618)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Chimeric BVDV/HCV NS3-wt sequence.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ABA95615 standard; DNA; 12734 BP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                99US-0344456
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           99US-0344456
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Lai VCH, Lau JYN;
                                                                                                                                                                                                                                                                                                                                                        1.28
49.00
100.00%
100.00%
100.00%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            21-MAR-2002 (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        (SCHE ) SCHERING CORP.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              WPI; 2002-121103/16.
          WPI; 1992-417213/51
                                                                                                                                                                                                                                                                                                                                                                                                Best Local Similarity:
                         P-PSDB; AAR29846
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        infection
                                                                                                                                                                                                                                                                                                                                                                                    Percent Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         US6326137-B1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                25-JUN-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           25-JUN-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    04-DEC-2001.
                                                                                                                                                                                                                                                                                                                                          Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Example 2;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Hong Z,
                                                                                                                                                                                                                                                                                                                                                                                                               Query Match:
DB:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        HCV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         RESULT 6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
```

```
chimeric Repatitis C virus (HCV)-pestivirus genome. The construct comprises a pestivirus genome where a Npro pestivirus protease gene is replaced with a gene encoding a functional HCV NS3 protease. Furthermore, each junction site recognised by the Npro protease is replaced with a junction site recognised by the HCV NS3 protease. The construct is useful for screening compounds that inhibit HCV in vivo by inhibiting HCV protease, where screening may be in cell culture or in an animal model. The present sequence is a chimmeric clone of BVDV (bovine viral diarrhea virus)/HCV NS3-wt, which was used to illustrate the present invention.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               The present invention describes a covalent hepatitis C virus (HCV) NS4A-NS3 complex comprising a central hydrophobic domain of native HCV NS4 serine protease domain, where the hydrophobic domain of native HCV NS4 septide is tethered by the linker to the amino terminus of the HCV NS3 protease domain. The covalent NS4A-NS3 complexes are useful for structural determination and determination of mode of binding of HCV inhibitors by NMR spectroscopy. They can also be used for detecting inhibitors of the protease activity, the helicase activity and the ATPase activity of NS3. The covalent
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                HCV: hepatitis C virus; single chain recombinant complex; linker;
NS3 protease; NS4A cofactor; NS4A-NS3 complex; serine protease;
hydrophobic domain; covalent complex; detection; inhibitor; primer; ss.
             present invention relates to a nucleic acid construct encoding a
                                                                                                                                                                                                         Seguence 12734 BP; 4032 A; 2604 C; 3295 G; 2803 T; 0 other;
                                                                                                                                                                                                                                                           12734
11
0
0
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    HCV NS4A-NS3 complex construction primer SEQ ID NO:48.
                                                                                                                                                                                                                                                                            Matches:
Conservative:
Mismatches:
                                                                                                                                                                                                                                                                                                                                                                                                                                          1 GlySerValVallleValGlyArglleValLeu 11
                                                                                                                                                                                                                                                                                                                                  Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              New hepatitis C virus covalent complexes
                                                                                                                                                                                                                                                                                                                                                                                      US-09-965-594-26 (1-11) x ABA95615 (1-12734)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Example 1; Page 32; 211pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Weber PC,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    980S-0094331
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  98WO-US24528
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    970S-0067315
                                                                                                                                                                                                                                                           39.4
49.00
100.008
100.008
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              AAX80310 standard; DNA; 75
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Malcolm BA, Taremi SS,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        (SCHE ) SCHERING CORP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           WPI; 1999-385385/32.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                 Best Local Similarity:
                                                                                                                                                                                                                                                                                                Percent Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                24-NOV-1998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  28-JUL-1998;
28-NOV-1997;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               W09928482-A2
                                                                                                                                                                                                                                              Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 07-SEP-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               10-JUN-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Synthetic
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 AAX80310;
                                                                                                                                                                                                                                                                                                                                    Query Match:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            RESULT 7
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             AAX80310
                                                                                                                                                                                                                                                                                 Score:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 οy
```

S

Page

```
The present invention describes a covalent hepatitis C virus (HCV) NS4A-NS3 complex comprising a central hydrophobic domain of native HCV NS3 serine protease domain, where the hydrophobic domain of native HCV NS3 serine protease domain, where the hydrophobic domain of native HCV NS3 protease domain. The covalent to the amino terminus of the HCV NS3 protease domain. The covalent NS4A-NS3 complexes are useful for structural determination and determination of mode of binding of HCV inhibitors by NMR spectroscopy. They can also be used for detecting inhibitors of the protease activity and the ATPase activity of NS3. The covalent NS4A-NS3 complexes are more soluble, stable and active than the non-covalent protease-peptide complexes previously available. The present sequence represents a primer used in the construction of the
                                                                                                                                                                                                                                                                                                                                                                                                                                                                        HCV; hepatitis C virus; single chain recombinant complex; linker; NS3 protease; NS4A cofactor; NS4A-NS3 complex; serine protease; hydrophobic domain; covalent complex; detection; inhibitor; primer; ss.
non-covalent protease-peptide complexes previously available. The present sequence represents a primer used in the construction of the HCV NS4A-NS3 complexes.
                                                                                                                                                                                                                                                                                                                                                                                                                                               HCV NS4A-NS3 complex construction primer SEQ ID NO:55.
                                                                                                                                  Length:
Matches:
Conservative:
Mismatches:
Indels:
Gaps:
                                                                                 Sequence 75 BP; 16 A; 14 C; 15 G; 30 T; 0 other;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Sequence 75 BP; 16 A; 12 C; 16 G; 31 T; 0 other;
                                                                                                                                                                                                                                                                                  GlySerValValIleValGlyArglleValLeu 11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             New hepatitis C virus covalent complexes
                                                                                                                                                                                                                                      US-09-965-594-26 (1-11) x AAX80310 (1-75)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Example 1; Page 33; 211pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Weber PC,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               98WO-US24528.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            98US-0094331
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        97US-0067315
                                                                                                                               0.186
48.00
100.00%
90.91%
97.96%
                                                                                                                                                                                                                                                                                                                                                         AAX80314 standard; DNA; 75
                                                                                                                                                                                                                                                                                                                                                                                                                   07-SEP-1999 (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Malcolm BA, Taremi SS,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        HCV NS4A-NS3 complexes.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (SCHE ) SCHERING CORP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               WPI; 1999-385385/32.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Synthetic.
Hepatitis C virus.
                                                                                                                                             Score:
Percent Similarity:
Best Local Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    W09928482-A2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            28-JUL-1998;
                                                                                                               Alignment Scores:
Pred. No.:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              24-NOV-1998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        28-NOV-1997;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                10-JUN-1999
                                                                                                                                                                                                                                                                                                                                                                                      AAX80314;
                                                                                                                                                                                                                                                                                              13
                                                                                                                                                                                        Query Match:
DB:
                                                                                                                                                                                                                                                                                                                         ប្រសួលស្ន
                                                                                                                                                                                                                                                                   င်
                                                                                                                                                                                                                                                                                              a
```

```
The present invention describes a covalent hepatitis C virus (HCV) NS4A NS3 complex comprising a central hydrophobic domain of native HCV NS4A peptide, a linker, and an HCV NS3 serine profease domain, where the hydrophobic domain of native HCV NS4A peptide is tethered by the linker to the amino terminus of the HCV NS3 protease domain. The covalent NS4A NS3 complexes are useful for structural determination and determination of mode of binding of HCV inhibitors by MNR spectroscopy. They can also be used for detecting inhibitors of the protease activity, the helicase activity and the ATPase activity of NS3. The covalent NS4A NS3 complexes are more soluble, stable and active than the non-covalent protease-peptide complexes previously available. The HCV NS4A-NS3 complexes.
                                                                                                                                                                                                                                                                                                                    HCV; hepatitis C virus; single chain recombinant complex; linker; NS3 protease; NS4A cofactor; NS4A-NS3 complex; serine protease; hydrophobic domain; covalent complex; detection; inhibitor; primer; ss.
                                                                                                                                                                                                                                                                                           HCV NS4A-NS3 complex construction primer SEQ ID NO:42.
           75
10
0
0
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       78
10
0
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Sequence 78 BP; 16 A; 14 C; 17 G; 31 T; 0 other;
                                    Conservative:
Mismatches:
                                                                                                                             Matches:
                                                               Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    New hepatitis C virus covalent complexes
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Yao
                                                                                                     US-09-965-594-26 (1-11) x AAX80314 (1-75)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Weber PC,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Example 1; Page 31; 211pp; English.
                                                                                                                                                                                                              AAX80306 standard; DNA; 78 BP.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                      98WO-US24528
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              98US-0094331.
         0.186
48.00
100.00%
90.91%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      0.195
48.00
100.00%
90.91%
97.96%
                                                                                                                                                                                                                                                                  07-SEP-1999 (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Malcolm BA, Taremi SS,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (SCHE ) SCHERING CORP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            WPI; 1999-385385/32.
                                                                                                                                                                                                                                                                                                                                                                           Synthetic.
Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Percent Similarity:
Best Local Similarity:
Query Match:
DB:
                                   Percent Similarity:
Best Local Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Alignment Scores:
Pred. No.:
                                                                                                                                                                                                                                                                                                                                                                                                                 W09928482-A2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                     24-NOV-1998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                28-JUL-1998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              28-NOV-1997;
Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                         0-JUN-1999
                                                                                                                                                                                                                                         AAX80306;
                                                               Query Match:
         Pred. No.:
                                                                                                                                                                                   RESULT 9
                                                                                                                                                                                                  AAX80306
                          Score:
                                                                                                                                                                                                                                         d
```

9

RESULT 10 AAX80290

```
HCV; hepatitis C virus; single chain recombinant complex; linker;
NS3 protease; NS4A cofactor; NS4A-NS3 complex; serine protease;
hydrophobic domain; covalent complex; detection; inhibitor; primer; ss.
                                                                                                                              HCV NS4A-NS3 complex construction primer SEQ ID NO:87.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           New hepatitis C virus covalent complexes
                                                                                                                                                                                                                                                                                                                                                                                                                                                            δ,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Example 2; Page 43; 211pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                             Weber
                              AAX80338 standard; DNA; 96 BP.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        AAX80340 standard; DNA; 96 BP
                                                                                                                                                                                                                                                                                                                                            98WO-US24528
                                                                                                                                                                                                                                                                                                                                                                         98US-0094331.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   100.00%
90.91%
97.96%
                                                                                                (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      0.246
                                                                                                                                                                                                                                                                                                                                                                                                                                                            Taremi SS,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       HCV NS4A-NS3 complexes.
                                                                                                                                                                                                                                                                                                                                                                                                                         (SCHE ) SCHERING CORP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         WPI; 1999-385385/32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Percent Similarity:
Best Local Similarity:
                                                                                                                                                                                                                              Synthetic.
Hepatitis C virus.
                                                                                                                                                                                                                                                                             W09928482-A2
                                                                                                                                                                                                                                                                                                                                           24-NOV-1998;
                                                                                                                                                                                                                                                                                                                                                                         28-JUL-1998;
28-NOV-1997;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Alignment Scores:
                                                                                                07-SEP-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                            Malcolm BA,
                                                                                                                                                                                                                                                                                                            10-JUN-1999
                                                               AAX80338;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Query Match:
RESULT 11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       RESULT 12
                 AAX80338
                                                                 qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   The present invention describes a covalent hepatitis C virus (HCV) NSAA-NS3 complex comprising a central hydrophobic domain of native HCV NSAA peptide. A linker, and an HCV NS3 serine protease domain, where the hydrophobic domain of native HCV NS4A peptide is tethered by the linker to the amino terminus of the HCV NS3 periodese domain. The covalent NS4A-NS3 complexes are useful for structural determination and determination of mode of binding of HCV inhibitors by NMR spectroscopy. They can also be used for detecting inhibitors by NMR spectroscopy. They can also be used for detecting inhibitors of the protease activity, the halicase activity and the Arpase activity of NS3 complexes are more soluble, stable and active than the
                                                                                                                                                                                                                                                                        HCV; hepatitis C virus; single chain recombinant complex; linker;
NS3 protease; NS4A cofactor; NS4A-NS3 complex; serine protease;
hydrophobic domain; covalent complex; detection; inhibitor; primer; ss.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   non-covalent protease-peptide complexes previously available. The present sequence represents a primer used in the construction of the HCV NS4A-NS3 complexes.
                                                                                                                                                                                                                                            HCV NS4A-NS3 complex construction primer SEQ ID NO:26.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      78
10
0
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Length:
Matches:
Conservative:
Mismatches:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Sequence 78 BP; 18 A; 11 C; 18 G; 31 T; 0 other;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1 GlySerValValIleValGlyArgIleValLeu 11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Yao
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      New hepatitis C virus covalent complexes
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   US-09-965-594-26 (1-11) x AAX80290 (1-78)
            JS-09-965-594-26 (1-11) x AAX80306 (1-78)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Example 1; Page 26; 211pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         PC,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Weber
                                                                                                                                                ВР
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       98US-0094331
97US-0067315
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 0.195
48.00
100.00%
90.91%
97.96%
                                                                                                                                              AAX80290 standard: DNA; 78
                                                                                                                                                                                                             (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Taremi SS,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (SCHE ) SCHERING CORP.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       WPI; 1999-385385/32
                                                                                                                                                                                                                                                                                                                                                       Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Percent Similarity:
                                                                                                                                                                                                                                                                                                                                                                                         W09928482-A2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       28-JUL-1998;
28-NOV-1997;
                                                                                                                                                                                                            07-SEP-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                        24-NOV-1998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                        10-JUN-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Malcolm BA,
                                                                                                                                                                                                                                                                                                                                         Synthetic.
                                                                                                                                                                           AAX80290;
```

Yao

```
The present invention describes a covalent hepatitis C virus (HCV) NS4A-NS1 complex comprising a central hydrophobic domain of native HCV NS4A peptide, a linker, and an HCV NS3 serine protease domain, where the hydrophobic domain of native HCV NS4A peptide is tethered by the linker to the amino terminus of the HCV NS4A peptide is tethered by the linker to the amino terminus of the HCV NS3 protease domain. The covalent determination and mode of binding of HCV inhibitors by NMR spectroscopy. They can also be used for detecting inhibitors of the protease activity, the helicase activity and the APPase activity of NS3. The covalent non-covalent protease Peptide complexes previously available. The present sequence represents a primer used in the construction of the
                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                               Seguence 96 BP; 21 A; 17 C; 21 G; 37 T; 0 other;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Matches:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      US-09-965-594-26 (1-11) x AAX80338 (1-96)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      07-SEP-1999 (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                AAXB0340
ID AAX8
XX AAX
AC AAX8
XX DT 07-S
```

13

Query Match: Best Local

Score:

Hepatitis C Virus

```
The present invention describes a covalent hepatitis C virus (HCV) NSAA NS3 complex comprising a central hydrophobic domain of native HCV NS4A peptide, a linker, and an HCV NS3 esrine protease domain, where the hydrophobic domain of native HCV NS4A peptide is tethered by the linker to the amino terminus of the HCV NS3A protease domain. The covalent NS4A-NS3 complexes are useful for structural determination and determination of mode of binding of HCV inhibitors by NMR spectroscopy. They can also be used for detecting inhibitors of the protease activity and the APPase activity of NS3. The covalent NS4A-NS3 complexes are more soluble, stable and active than the non-covalent protease-peptide complexes previously available. The present sequence represents a primer used in the construction of the
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Double stranded RNA; dsRNA; RNAI; RNA inhibition; cytostatic; virucide; proteozoadide; gene expression; antisense; tumour; infection; Plasmodium; virus; viroid; anti-GFP; human; HIV; human immunodeficiency virus; Hepatitis C virus; human papalloma virus; gene; ds.
                                                 HCV; hepatitis C virus; single chain recombinant complex; linker;
NS3 protease; NS4A cofactor; NS4A-NS3 complex; serine protease;
hydrophobic domain; covalent complex; detection; inhibitor; primer; ss.
                 HCV NS4A-NS3 complex construction primer SEQ ID NO:89
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Hepatitis C virus NS4A polynucleotide SEQ ID NO 129.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               96
10
0
0
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Length:
Matches:
Conservative:
Mismatches:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Sequence 96 BP; 21 A; 17 C; 21 G; 37 T; 0 other;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 31 GGTTCTGTTGTTATTGTTGCTAGAATTATTTTA 63
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1 GlyServalValIleValGlyArglleValLeu 11
                                                                                                                                                                                                                                                                                                                                                          Weber PC, Yao N;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                          New hepatitis C virus covalent complexes
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 US-09-965-594-26 (1-11) x AAX80340 (1-96)
                                                                                                                                                                                                                                                                                                                                                                                                                                                        Example 2; Page 44; 211pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ABZ35821 standard; DNA; 161 BP
                                                                                                                                                                                                                                     98WO-US24528.
                                                                                                                                                                                                                                                                     980S-0094331.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              0.246
48.00
100.00%
90.91%
97.96%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      07-FEB-2003 (first entry)
                                                                                                                                                                                                                                                                                                                                                          Malcolm BA, Taremi SS,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               HCV NS4A-NS3 complexes.
                                                                                                                                                                                                                                                                                                                        (SCHE ) SCHERING CORP
                                                                                                                                                                                                                                                                                                                                                                                       WPI; 1999-385385/32.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Score:
Percent Similarity:
                                                                                                                                   Hepatitis C virus
                                                                                                                                                                                                                                                                   28-JUL-1998;
28-NOV-1997;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Alignment Scores:
Pred. No.:
                                                                                                                                                                                                                                     24-NOV-1998;
                                                                                                                                                                     W09928482-A2
                                                                                                                                                                                                     10-JUN-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      AB235821;
                                                                                                                     Synthetic
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   RESULT 13
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     AB235821
IID AB2
IXX
AC AB2
XX
XX
XX
XX
DDT 07-
DX
XX
DE Hep
XX
XX
DE Hep
XX
XX
XX
MW DOU
셤
```

```
The invention relates to inhibiting expression of a target gene in a cell by introducing at least two oligoribonucleotides (dsrNAI and II), both with a double-stranded (ds) structure of at most 49 sequential nucleotide pairs. At least part of one strand (sl., $2) of the ds structures in each of dsrNAI and II are complementary to regions in the target gene. The method uses antisense inhibition of gene expression using double stranded RNA inhibition (RNAI). The method is particularly used to treat tumours or infections, especially by plasmodium or viruses/viroids (pathogenic on humans, animals or plants). The method provides more effective inhibition of expression than known methods using a single dsRNA, even at very low concentrations. When dsRNA has at least one unpaired nucleotide at the end, stability (and thus effective concentration in the cells improved and efficiency can be increased further by pretreating the cells with interferon. The present sequence is that of a target DNA of the
                                                                                                                                                                                                                                                             Inhibiting expression of target genes, useful e.g. for treating tumors, by introducing into cells two double-stranded RNAs that are complementary to the target
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 interferon; oncogene; cytokine; Id; developmental;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     161
0 0 0
0 0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       55 G; 34 T; 0 other;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Conservative:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ä
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Length:
Matches:
                                                                                                                                                                                                  Hadwiger
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            US-09-965-594-26 (1-11) x ABZ35821 (1-161)
                                                                                                                                                                                                                                                                                                                           Claim 13; Page 87; 100pp; German.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  HCV NS4A DNA fragment SEQ ID 129
                                                                                                                                                                                                  Rost S,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ABX10064 standard; DNA; 161 BP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Sequence 161 BP; 32 A; 40 C;
                                                                                                      09-JAN-2001; 2001DE-1000588.
                                                                                                                                    09-JAN-2001; 2001DE-1000588
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   0.442
48.00
100.00%
90.91%
97.96%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 23-JAN-2003 (first entry)
                                                                                                                                                                                                  Kreutzer R, Limmer S,
                                                                                                                                                                   (RIBO-) RIBOPHARMA AG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Oligoribonucleotide;
prion; inhibition; ds
                                                                                                                                                                                                                               WPI; 2002-683450/74.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Percent Similarity:
Best Local Similarity:
Query Match:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Hepatitis c virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            DE10100587-C1.
                                            DE10100588-A1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Alignment Scores:
                                                                           18-JUL-2002
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           invention
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ABX10064;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      RESULT 14
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ABX10064
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Score:
```

21-NOV-2002

œ

07-DEC-2001; 2001DE-1060151.

```
Search completed: August 30, 2003, 19:48:23 Job time : 12.2149 secs
  õ
                                                                                                                                                                                                                          This invention describes a novel method for inhibiting expression of a target gene by introducing into the cell that contains the target gene at least one oligoribonuclectide (GRNAI) that has a double-stranded (GAS) structure of not more than 49 consecutive nucleotides (nt), where at least a segment of one strand of the ds structure is complementary with the target gene and the cells are treated with interferon before introduction of dsRNAI. The method is used to inhibit expression of target genes, particularly oncogenes, cytckine genes, Id (not defined) protein genes; developmental or prior genes expressed in pathogenic organisms (particularly plasmodia) or in viruses or virois (pathogenic in humans, animals or plants). Treating the cells with interferon greatly increases the extent to which dsRNA can inhibit expression of the target genes, and the effect is even greater when dsRNA are modified to increase their stability, ABX09936-ABX10075 represent
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        RNA inhibition; dsRNA1; gene expression inhibitor; oncogene; cytostatic; virucide; protozoacide; gene; ds.
                                                                                                                                                Inhibiting expression of target genes, e.g. oncogencs, in cells, tintroduction of complementary double-stranded oligoribonucleotide, after treating the cell with interferon
                                                                                                                                                                                                                                                                                                                                                                                                                       gene fragments used to illustrate the method of the invention
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1000
                                                                                                                                                                                                                                                                                                                                                                                                                                                   Sequence 161 BP; 32 A; 40 C; 55 G; 34 T; 0 other;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Length:
Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1 GlySerValValileValGlyArgileValLeu 11
                                                                                             Hadwiger
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Hepatitis C virus NS4A DNA SEQ ID NO 129
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      US-09-965-594-26 (1-11) x ABX10064 (1-161)
                                                                                                                                                                                                       Disclosure; Page 92; 98pp; German
                                                                                            ROSt S,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 BP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           09-JAN-2001; 2001DE-1000586.
26-OCT-2001; 2001DE-1055280.
29-NOV-2001; 2001DE-1058411.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 09-JAN-2002; 2002WO-EP00152
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        0.442
48.00
100.00%
90.91%
97.96%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ABV78245 standard; DNA; 161
             09-JAN-2001; 2001DE-1000587
                                        09-JAN-2001; 2001DE-1000587
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    15-NOV-2002 (first entry)
                                                                                            Kreutzer R, Limmer S,
                                                                   (RIBO-) RIBOPHARMA AG
                                                                                                                       WPI; 2002-742209/81
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Best Local Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Alignment Scores:
Pred. No.:
Score:
Percent Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            WO200255693-A2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       18-JUL-2002
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ABV78245;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            09
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Query Match:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       RESULT 15
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ABV78245
q
```

```
The invention relates to inhibiting expression of a target gene (1) in a cell by introducing an inhibitory RNA (dsRNA1) having a double-stranded structure of at most 49 consecutive bases. At least part of one strand (as1) of dsRNA1 is complementary to (1) and at least part of one strand has an overhang of 1-4 nucleotides. The method is used to inhibit the expression of a wide range of genes, e.g. oncogenes, cytokine genes etc. in humans, also genes in Plasamodium or in viruses or viroids that are pathogenic for humans, animals or plants. Introducing an overhang into dsRNA greatly increases effectiveness for inhibiting gene expression, both in vivo and in vitro and also increases stability and thus the effective concentration inside the cell. The present sequence is that of
                                                                                                                                                                                           the
                                                                                                                                                             expression of target gene, useful e.g. for inhibiting by administering double-stranded RNA complementary to
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      161
10
10
0
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Sequence 161 BP; 32 A; 40 C; 55 G; 34 T; 0 other;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   60 GGCAGCGTGGTCATTGTGGCAGGATCATCTTG 92
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1 GlySerValVallleValGlyArglleValLeu 11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Matches:
                                                       Hadwiger
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Length:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              US-09-965-594-26 (1-11) x ABV78245 (1-161)
                                                                                                                                                                                                                                                                       Claim 10; Page 190; 203pp; German.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   gene related to the invention.
                                                    Rost S,
                                                                                                                                                                                           oncogenes, by administering do
target and having an overhang
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      0.442
48.00
100.00%
90.91%
97.96%
                                                    ŝ
(RIBO-) RIBOPHARMA AG
                                                                                                         WPI; 2002-590671/63
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Percent Similarity:
Best Local Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Alignment Scores:
                                                       Kreutzer R,
                                                                                                                                                          Inhibiting
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Query Match:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ..
Q
```

GenCore version 5.1.6 Copyright (c) 1993 - 2003 Compugen Ltd.

```
August 30, 2003, 19:20:43; Search time 106.667 Seconds (without alignments) 2506.388 Million cell updates/sec
OM protein - nucleic search, using frame_plus_p2n model
                                                                                                                                                                                                                                                             22781392 seqs, 12152238056 residues
                                                                                                                                                                                                                                                                                           hits satisfying chosen parameters:
                                                                                                                                                                              0.5
                                                                                                                                                         BLOSUM62
Xgapop 10.0 , Xgapext
Ygapop 10.0 , Ygapext
Fgapop 6.0 , Fgapext
Delop 6.0 , Delext
                                                                                              US-09-965-594-26
49
1 GSVVIVGRIVL 11
                                                                                               Title:
Perfect score:
Sequence:
                                                                                                                                                                Scoring table:
                                                                                                                                                                                                                                                              Searched:
                                Run on:
```

Post-processing: Minimum Match 0% Maximum Match 100% Listing first 45 summaries Minimum DB seq length: 0
Maximum DB seq length: 2000000000

oţ

rotal number

45562784

EST:\* Database

gb\_est3:\*
gb\_est4:\*
gb\_est5:\*
em\_estfun:\*
em\_estom:\*
em\_gss\_hum:\* em\_gss\_mus:\*
em\_gss\_rod:\*
em\_gss\_rod:\*
em\_gss\_vrl:\*
gb\_gssl:\* em\_gss\_fun:\* em\_gss\_mam:\* em\_gss\_inv:\* em\_gss\_pln:\* em\_gss\_vrt:\* em\_estba:\* em\_esthum:\* em\_estin:\* cm\_estmu:\* em\_estov:\* em\_estpl:\* em\_estro:\* em\_htc:\* gb\_est2:\* gb\_htc:\* gb\_est1:\* 

29: gb\_gss2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

ARIES		Description	5 AWSB0055 F	53 CA728563	B7 BW300987	38 BW306038	BW176544	TRECORDER PERCESSES	#00000000 70000000000000000000000000000	07 BQ1210/3	39 81718639	6 PER 10	63 BI882963	09 BE556909	86 CB417286	53 BI981053	7/R04918 7/	51 BOOG3551	02 BI982002	43 BE016243	38 BE016238	05 BE201405	66 CD3 67 22 25 3 9 2 2 2 3 9 2 2 3 9 3 9 2 2 3 9 9 9 9	AW019436 1	0 PT 77 TV	35 BI759235	93 CC250893	AV211292 P	OL AL264414	54 BG#03954	592 CB030392 IGESTZY 158 B2004158 OGG8040	63 BO785063	61 AG096061	09 AG092509	60 BG708760	52 BC035952	69 BE069569	BE069567	33 BEO/1133	30 BE069730	DECITION RC3-	77047 \$6771044
SUMMARI		a	W5800	728	300	90	9 7 9	200	,,,	18	718	333	B1882	BE556	CB417	B1981	D 1040	BO093	BI982	BE016	BE016	BE201	CD392	2,5	75.7	759	250	AV2112	CNSO3	BC404	B20041	BO785	AG096	AG092	BG708	BC035	BE069	BEUDY	BEOLE	20000	BV0122	
		80	Q	14	13		7 :	3 6	1 -	10	3 .	b	12	10	7	12	7.5	13	12	10	10	9	7 (	ס ת	٦, ر	12	29	σ	52	7.	1 6	13	29	29	. 10	Ξ:	10	3 .	7 .	7 .	3 0	٠.
		Length	•	~ 1	ø.	0 0	V	0 =	r w	) r	m	22	8	0	S	<b>8</b> 0 0	n c	0	N	C	*	<b>*</b> (	o t	~ ₫	Λ 00	8	9	-	ъ.	or us	200	-	79	28	ŝ	6	~ <	2 6	٠.	<b>~</b> 0	. ~	3 4
d	Query	Match	Ξ.							 	Ś	۳.	٠. ن	٠. ش	m.	m c			ω.	ω.	٠ س	m (	n c			m.	Э.	ä,	∹.	∹ ←	81.6	; ;	;	ä	<u>.</u>	·	٠ د	n c	n c	n o	n o	. a
		Score	45	4.	43			4 4				41						4.1						4.4			41	0 9	7	• •	<b>4</b>	4	40	0.4	40	40	ж Ж	ν c	900	, c	n 0	0 0
	Result	Q		o 0	<b>.</b>		n u	۰ ۲	- α	0	10	11	c 12	13	٦,	2 15	- ۲	1 ~	М	N	~	<b>(1)</b>	71 (	2 C	1 ~	10	c 28	~ 1	ک ک د	3.7	4 (F)	(17)	c 35	m ·	37	80 c	ν • ν •	<b>7</b> ₹	<b>+</b> C	7 7	4	

ALIGNMENTS

AM580055 234 bp mRNA linear EST 16-MAR-2000 CSC19T0355-130100-011-905 HT0375 Homo sapiens cDNA, mRNA sequence. AM580055.1 GI:7255104 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. 1 (bases 1 to 234)
HCGP http://www.ludwig.org.br/ORESTES. Homo sapiens (human) Homo sapiens EST. LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM REFERENCE AUTHORS RESULT 1 AW580055

~

us-09-965-594-26.rst

```
cocation/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          259
43.00
100.008
90.008
87.768
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        145 c
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             235
43.00
90.91%
72.73%
87.76%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Ciona intestinalis
Ciona intestinalis
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    complex.
                Seq primer: M13.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Percent Similarity:
Best Local Similarity:
Query Match:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Percent Similarity:
Best Local Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ø
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        201
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Aliqnment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Query Match:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                BASE COUNT
ORIGIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Pred. No.:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ..
No.:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       DEFINITION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            BASE COUNT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              BW300987
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          RESULT 3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         FEATURES
                                                    FEATURES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ORIGIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1...234
//organism="know sapiens"
//organism="know sapiens"
//mol_type="mkNa"
//dev_stage="mkNa"
//dev_stage="mkNa"
//dev_stage="mkNa"
//dev_stage="mkn375"
//dev_stage="mkn375"
//dev_stage="mkn375"
//dev_stage="mkn375"
//dev_stage="mkn375"
//dev_stage="mkn3"
//dev_stage="mkna"
/
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                wdilc.pk004.f22 wdilc Triticum aestivum cDNA clone wdilc.pk004.f22 5' end, mRNA sequence. CA728563 GA728563 GEST CA728563 EST.
                                                                                                                                                                                                                                                          Fax: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl-RCl&t2-RCl-HT0375-
130100-011-905&t3-2000-01-13&t4-1)
Seq primer: puc l@ forward
High quality sequence start: 28
High quality sequence stop: 216.
Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Bukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta: Liliopsida; Poales; Poaceae; Pooideae; Triticeae; Tr
                                        Unpublished
Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          E. I. DuPont de Nemours and Company
1 Innovation Way, P.O. Box 6104, Newark, DE 19714-6104, USA
1 Innovation Way, P.O. Box 6104, Newark, DE 19714-6104, USA
1 Innovation Way, P.O. Box 6104, Newark, DE 19714-6104, USA
Fax: 302-631-2602
Email: Scott.V.Tingey@USA.dupont.com
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           23.4
0 0 0 0
        FAPESP/LICR Human Cancer Genome Project
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Length:
Matches:
Conservative:
Mismatches:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Triticum aestivum (bread wheat)
Triticum aestivum
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              US-09-965-594-26 (1-11) x AW580055 (1-234)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Contact: Scott V. Tingey
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           26.4
45.00
100.00%
72.73%
91.84%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Crop Genetics
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Percent Similarity:
Best Local Similarity:
Query Match:
DB:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Unpublished
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   RESULT 2
CA728563/c
LOCUS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      BASE COUNT
ORIGIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    DEFINITION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 TITLE
JOURNAL
COMMENT
TITLE
JOURNAL
COMMENT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                AUTHORS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        REFERENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Score:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          à
```

```
BW300987 Nori Satch unpublished cDNA library, neural complex Ciona intestinalis cDNA clone cinc0209065', mRNA sequence.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Eukaryota, Metazoa, Chordata, Urochordata; Ascidiacea, Enterogona, Phlebobranchia; Cionidae; Ciona.
I (bases I to 664)
Satou, Y., Shin i, T., Kohara, Y. and Satoh, N.
Expressed genes in Ciona intestinalis (2002c)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             /tissue_type="neural complex" /clone_lib="Nori Satoh unpublished cDNA library, neural
                                                                                                                                                                                                                                                                                                Site_1: EcoRI; Site_2:
Hi Line) developing
                                                                                                                                                                                                                                                                                                                                                                                                                          6 others
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             618
2
1
0
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              6664
0 0 0 0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Unpublished
Contact: Nort Satch
Contact: Nort Satch
Experient of Zoology
Kyoto University
Sakyo-ku, Kyoto, Kyoto 606-8502, Japan
Tel: 81-75-753-4081
Fax: 81-75-705-1113
Email: satch@ascidian zool.kyoto-u.ac.jp.
                      /usgailsmin_illicum_destivum
/mol_type="mRNA"
/db_xref="taxon:4565"
/clone="wdilc.pk004.f22"
/tissue_type="inflorescence"
/lab_host="billorescence"
/clone_lib="wdilc"
/note="wector: pBluescript SK+; Si
/note="wector: pBluescri
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Length:
Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Length:
Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  /organism="Ciona intestinalis"
/mol_type="mkna"
/db_xref="taxon:7719"
/clone="cinc020906"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      /organism-*Triticum aestivum*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            183 t
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Gaps:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              US-09-965-594-26 (1-11) x CA728563 (1-618)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       135 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            BW300987
BW300987.1 GI:24881598
```

```
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl-&t2-RC2-BT0389-120
400-014-c02&t3-2000-04-12&t4-1)
Seq primer: puc 18 forward
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    BE069681 150400-014-c02 BT0389 Homo sapiens cDNA, mRNA sequence.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. 1 (bases 1 to 150) Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R., Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F., Goldman, G.H., Carvalho, A.F., Matsukuma, A., Baia, G.S., Simpson, D.H., Brunstein, A., deoliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and
                                                                                    Shotgun sequencing of the human transcriptome with ORF expressed sequence tags Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000) 20202663
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
                                                        Email: satoh@ascidian.zool.kyoto-u.ac.jp.
                                                                                                                                                                                                                                                                                                                Conservative:
Mismatches:
       Sakyo-ku, Kyoto, Kyoto 606-8502, Japan
Tel: 81-75-753-4081
Fax: 81-75-705-1113
                                                                                                                                                                                                                                                                               Length:
Matches:
                                                                                                                                                                                                                                                                                                                                                                                                                               Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      /organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/dev_stage="Adult"
                                                                                                                                                                                                                                                                                                                                                                    Gaps:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   High quality sequence start: 3
High quality sequence stop: 150.
Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                US-09-965-594-26 (1-11) x BW176544 (1-725)
                                                                           Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    BE069681
BE069681.1 GI:8414331
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Homo sapiens (human)
                                                                                                                                                                                                                                                                                             43.00
100.00%
90.00%
87.76%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Homo sapiens
                                                                                                                                                                                                                                                                                                        Percent Similarity:
Best Local Similarity:
Query Match:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              10737800
                                                                                                                                                                                                              195
                                                                                                                                                                                                                                                             Alignment Scores:
Pred. No.:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           source
                                                                                                                                                                                                          BASE COUNT
ORIGIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    DEFINITION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ORGANISM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         JOURNAL
MEDLINE
PUBMED
COMMENT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ACCESSION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  RESULT 6
BE069681
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     KEYWORDS
SOURCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      REFERENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        AUTHORS
                                                                        FEATURES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        FEATURES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         TITLE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     VERSION
                                                                                                                                                                                                                                                                                                 Score:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                 qq
                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                    BW306038 Nori Satoh unpublished cDNA library, heart Clona intestinalis cDNA clone ciht012f15 5', mRNA sequence.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      BW176544 Nori Satoh unpublished cDNA library, heart Ciona intestinalis cDNA clone rciht012f15 3', mRNA sequence.
                                                                                                                                                                                                                                                                    Eukaryota, Metazoa, Chordata, Urochordata, Ascidiacea, Enterogona, Phlebobranchia, Cionidae, Ciona.
I (bases 1 to 709)
Satou, Y., Shin i, T., Kohara, Y. and Satoh, N.
Expressed genes in Ciona intestinalis (2002c)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Eukaryota: Metazoa; Chordata; Urochordata; Ascidiacea; Enterogona; Phlebobranchia; Cionidae; Ciona.

1 (bases 1 to 725)
Satou, Y., Shin-i, T., Kohara, Y. and Satoh, N.
Expressed genes in Ciona intestinalis (2002c)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           709
1
0
0
                                                                                                                                                                                                                                                                                                                                                                       Contact: Nori Satoh
Department of Zoology
Kyoto University
Sakyo-ku, Kyoto, Kyoto 606-8502, Japan
Tel: 81-75-753-4081
Fax: 81-75-705-1113
Fmail: Satoh@ascidian.zool.kyoto-u.ac.jp.
Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Length:
Matches:
Conservative:
Mismatches:
                                                 418 GGAAGTGGTGTTGGTTGAAGGATTGTT 447
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        307 GGAAGTGTGGTGGTTGGAAGGATTGTT 336
                                      10
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1 GlySerValValIleValGlyArgIleVal 10
                               1 GlySerValValIleValGlyArgIleVal
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   US-09-965-594-26 (1-11) x BW306038 (1-709)
JS-09-965-594-26 (1-11) x BW300987 (1-664)
                                                                                                                                                                                       BW306038.1 GI:24886649
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         BW176544.
BW176544.1 GI:24566468
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Contact: Nori Satoh
Department of Zoology
Kyoto University
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 283
43.00
100.00%
90.00%
87.76%
                                                                                                                                                                                                                                        Ciona intestinalis
Ciona intestinalis
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Ciona intestinalis
Clona intestinalis
                                                                                                                                                                                                                                                                                                                                                             Unpublished
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Percent Similarity:
Best Local Similarity:
Query Match:
DB:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              EST.
                                                                                                                                                                                                                        EST.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      LOCUS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         VERSION
KEYWORDS
SOURCE
ORGANISM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      RESULT 5
BW176544/c
                                                                                                                                                     DEFINITION
                                                                                                                                                                                                                                                        ORGANISM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                BASE COUNT
ORIGIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Pred. No.:
                                                                                                                                                                                                                                                                                                                                       TITLE
JOURNAL
COMMENT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            REFERENCE
AUTHORS
TITLE
                                                                                                                                                                                       ACCESSION
                                                                                                                                                                                                       VERSION
KEYWORDS
SOURCE
                                                                                                                                                                                                                                                                                                                           AUTHORS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               JOURNAL
                                                                                                                                                                                                                                                                                                         REFERENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ACCESSION
                                                                                                  RESULT 4
BW306038
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               FEATURES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Score:
                                                                                                                                       LOCUS
```

```
EST606651 mixed potato tissues Solanum tuberosum cDNA clone STMEV29 5' end, mRNA sequence.
B0121075.2 GI:21920306
EST.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            /clone="STMCY29"
/tissue_type="mixed tissues"
/lab_host="sol"mixed potato tissues"
/clone_lib="mixed potato tissues"
/clone_lib="mixed potato tissues"
/clone='vector: pBluescript SK(-); Site_l: EcoRI; Site_2:
XhoI; supplier: Combination of untreated and Phytophthora infestens-treated libraries of stolons, leaves, leaflets, axillary buds of stem explants, petioles, germinating eyes
      this sequence is actually derived from Phytophthora rather than potato." 110~c~163~g~214~t
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       on Apr. 17, 2002 this sequence version replaced gi:20173037. Contact: Robin Buell
The Institute for Genomic Research
9712 Medical Center Dr. Rockville, MD 20850, USA
Email: potato-array@tigr.org
This clone can be obtained from the University of Arizona Genomics
Institute. Orders can be made through URL:
http://genome.arizona.edu/orders/
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Solanum tuberosum (potato)
Solanum tuberosum (potato)
Solanum tuberosum
Solanum tuberosum
Solanum tuberosum
Solanum tuberosum
Spermatophyta; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnollophyta; eudicotyledons; core eudicots;
Saterids; Damidas; Solanales; Solanaceae; Solanum.

1 (bases 1 to 658)
Buell,C.R., Hart.A., Baker,B., Tanksley,S., Fry,W., Smart,C.,
Buell,C.R., Hartifiths,H., van der Hoeven,R., Tsai,J. and
Karamycheva,S.A.
Generation of a set of potato CDNA clones for microarray analyses
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    658
1
1
0
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Location/qualifiers
1. 658
/organism-"Solanum tuberosum"
/mol_type-"mRNA"
/cultivar-"Kennebec or Binjte"
/db_xref="taxon:4113"
                                                                                                                                                                                    Length:
Matches:
Conservative:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Conservative: Mismatches:
                                                                                                                                                                                                                                                                                                                                                                                                                                           Mismatches:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   220 t
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Matches:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Indels:
                                                                                                                                                                                                                                                                                               Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Length:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Gaps:
                                                                                                                                                                                                                                                                                                                                                                                       US-09-965-594-26 (1-11) x BG590884 (1-646)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       tubers, or roots."
97 c 182 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            42.00
90.91%
81.82%
85.71%
                                                                                                                                                                                                      42.00
90.91%
81.82%
85.71%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Seq primer: T3.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Unpublished
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Percent Similarity:
Best Local Similarity:
                                                                                                                                                                                                                                                                         Best Local Similarity:
                                                                   159 a
                                                                                                                                                                                                                                           Percent Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Alignment Scores:
                                                                                                                                                     Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Query Match:
DB:
                                                                                                                                                                                                                                                                                                  Query Match:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                BASE COUNT
ORIGIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       VERSION
KEYWORDS
SOURCE
ORGANISM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Pred. No.:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  DEFINITION
                                                                   BASE COUNT
ORIGIN
                                                                                                                                                                                       . . ON .
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      TITLE
JOURNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ACCESSION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        REFERENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   AUTHORS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            RESULT 8
BQ121075
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             FEATURES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                COMMENT
                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 l. 646
/organism="Solanum tuberosum"
//mol_type="mRNA"
/cultivar="texnon:4113"
/clone="BriffH5"
/fissue_type="leaf"
/dev_stage="6 week old"
/dev_stage="6 week old"
/dev_stage="6 week old"
/dev_stage="6 week old"
/dev_stage="leaf"
/dev_stage="leaf
/clone_lib="BT0389"
/note="Organ: breast; Vector: pucl8; Site_l: SmaI; Site_2: SmaI; Amin:-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196 ,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           BG590884 646 bp mRNA linear EST 07-MAR-2003 EST498726 P. infestans-challenged leaf Solanum tuberosum cDNA clone BPLI6H5 5' sequence, mRNA sequence.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   [ (bases 1 to 646) Zhang, M., Tornqvist, C.-E. Wirtz, U., Loukoianov, A., Zhang, P., Harberlach, G.T., Cho, J., Chiemingo, A., Bougri, O., Buell, C.R., Ronning, C.M., Helgeson, J. and Baker, B. Generation of ESTs from Potato Leaves Challenged with Phytophthora infestans, Incompatible Reaction
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         The Institute for Genomic Research 9712 Medical Center Dr. Rockville, MD 20850, USA Bmail: Potato-array@tigr.org
This clone can be obtained from the University of Arizona Genomics Institute. Orders can be made through URL:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Solanum tuberosum (potato)
Solanum tuberosum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotylodons; core eudicots;
asterids; lamiids; Solanales; Solanaceae; Solanum.
                                                                                                                                                                                                                                                                                                                                                      Length:
Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1 GlySerValValIleValGlyArgIleValLeu 11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        http://genome.arizona.edu/orders/
Seg primer: M13F-R.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Gaps:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               US-09-965-594-26 (1-11) x BE069681 (1-150)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 BG590884.1 GI:13609024
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Unpublished
Contact: Robin Buell
                                                                                                                                                                                                                                                                                                                                                         55.3
42.00
90.91%
72.73%
85.71%
                                                                                                                                                                                                                                                                                                                                                                                                             Percent Similarity:
Best Local Similarity:
Query Match:
                                                                                                                                                                                                                                                                                                                                   Alignment Scores:
Pred. No.:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          source
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 LOCUS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ACCESSION
VERSION
KEYWORDS
SOURCE
                                                                                                                                                                                                                                           BASE COUNT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ORGANISM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                REFERENCE
AUTHORS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       RESULT 7
BG590884
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        FEATURES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              TITLE
```

a à

J

S

```
AU183349.1 GI:13919090
                                                                                                                                                                                                                                                                                         Email: hanzg@chgc.sh.cn
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Contact: Masahiro Sakai
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Faculty of Agriculture
        Schistosoma japonicum
                    Schistosoma japonicum
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      42.00
90.91%
72.73%
85.71%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Best Local Similarity:
Query Match:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            11742760
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           21607774
                                                                                                                                                                                                                                                                                                                                                                                                                                                                      201
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Percent Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Alignment Scores:
                                                                                                                                                                                                                                                                                                                          source
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       DEFINITION
                    ORGANISM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Pred. No.:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ORGANISM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                      BASE COUNT
                                                                                                                                                    TITLE
JOURNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       REFERENCE
AUTHORS
TITLE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         MEDLINE
PUBMED
COMMENT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      RESULT 11
AU183349
                                                                      REFERENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ACCESSION
                                                                                      AUTHORS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          JOURNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       VERSION
KEYWORDS
                                                                                                                                                                                                                                                                                                         FEATURES
                                                                                                                                                                                       COMMENT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ORIGIN
      SOURCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SOURCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  /note="Vector: pBluescript SK(-); Sitc_1: EcoR1: Site_2: XhoI: Various sizes of sprouting eyes (2mm to 15mm) were taken from tubers. The tubers were incubated at 26C in the dark for 2-3 weeks prior to sprouting. The eyes were frozen in liquid nitrogen immediately upon removal from tubers."
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     BU718639 10-FEB-2003 SJM Schistosoma japonicum cDNA, mRNA sequence. BU718639 GI:28325986 EST.
                                                                                                                                 BG600305 777 bp mRNA linear EST 07-MAR-2003 EST505200 cSTS Solanum tuberosum cDNA clone cSTS28L13 5' sequence,
                                                                                                                                                                                                                                              Solanum tuberosum
Eukaryota, Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Wagnoliophyta; eudicotyledons; core eudicots;
asterids; lamids; Solanales; Solanaceae; Solanum.
1 (bases 1 to 777)
van der Hoeven, R., Bezzerides, J., Sun, H., Cho, J., Chiemingo, A.,
Bougri, O., Buell, C.R., Ronning, C., Tanksley, S. and Baker, B.
                                                                                                                                                                                                                                                                                                                                                                                  Unpublished
Contact: Robin Buell
The Institute for Genomic Research
9712 Medical Center Dr. Rockville, MD 20850, USA
Email: potato-arrayetigr.org
This clone can be obtained from the University of Arizona Genomics
Institute. Orders can be made through URL:
http://genome.arizona.edu/orders/
Seq primer: M13F-R.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               /db_xref-"taxon:4113"
/clone="cSTS28L13"
/clone="cSTS28L13"
/tissue_type="sprouting eyes from tubers"
/dev_stage="l2-14 weeks post harvest"
/iab_host="SOLR"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Length:
Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  /organism-"Solanum tuberosum"
/mol_type-"mRNA"
/cultivar-"Kennebec"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   US-09-965-594-26 (1-11) x BG600305 (1-777)
US-09-965-594-26 (1-11) x RQ121075 (1-658)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   196 9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  /clone_lib="cSTS"
                                                                                                                                                                                                                                    Solanum tuberosum (potato)
                                                                                                                                                                                              BG600305.1 GI:13617441
EST.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 132 C
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   499
42.00
90.91%
81.82%
85.71%
                                                                                                                                                                mRNA sequence.
BG600305
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Best Local Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Percent Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Query Match:
DB:
                                                                                                                                                                                                 VERSION
KEYWORDS
SOURCE
ORGANISM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      source
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   DEFINITION
ACCESSION
VERSION
                                                                                                                                                 DEFINITION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Pred. No.:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   BASE COUNT
                                                                                                                                                                                                                                                                                                                                                                TITLE
JOURNAL
COMMENT
                                                                                                                                                                                                                                                                                                                   REFERENCE
AUTHORS
                                                                                                                                                                                  ACCESSION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      RESULT 10
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     BU718639
LOCUS
                                                                                                RESULT 9
BG600305
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        KEYWORDS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   FEATURES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ORIGIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Score:
                                                                                                                                   LOCUS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
```

```
Eukaryota: Metazoa: Platyhelminthes: Trematoda: Digenea: Strigeidida: Schistosomatoidea: Schistosomatoidea: Schistosomatoidea: Schistosomatoidea: Liuis. Schistosomatoidea: Liuis. Schistosomatoidea: Liuis. Xuix., Zhuiz., Zhangix., Wangiy., Wangiy., Wangiz., Zhangix., Wangiy., Wangiy., Wangiy., Wangiy., Wangiy., Xuiz., Huangiy., Wangis., Wangiz., Xue., Xuiz., Huangiy., Wangis., Wangiz., Xue., Cheniz. and Han
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Dukaryota: Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes; Cyprinidea; Cyprinidea; Cyprinides; Cyprinidea; Cyprinides; Cyp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  EST 30-JUL-2002
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          AU183349 222 bp mRNA linear EST 30-JUL-200 AU183349 Cyprinus carpio head kidney stimulated by lipo-polysaccharide and concanavalin-A Cyprinus carpio cDNA clone H7, mRNA sequence.
                                                                                                                                                                                                                                                                                                                                                                    ^{\prime}2.
Expressed sequence tags from male adults of Schistosoma japonicum
                                                                                                                                                                                                                                                                                                                                                                                                                                                       Unpublished
Contact: Zeguang Han
Contact: Zeguang Han
Chinese National Human Genome Center at Shanghai
351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai
201203, P. R. China
Tel: 86-21-50801919(ex.45)
Fax: 86-21-50801922
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Miyazaki University
1-i nishi gakuenkibanadai, Miyazaki, Miyazaki 889-2192, Japan
Email: m.sakalecc.miyazaki-u.ac.jp.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    /organism="Schistosoma japonicum"
/mol_type="mRNA"
/db_xref="taxon:6182"
/sex="male"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Conservative:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           975 GCCAGTATAGTTGCCGTGGGTCGGGTCGTACTG 1007
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1 GlyServalVallleValGlyArglleValLeu 11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Mismatches:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Length:
Matches:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      /tissue_type="Whole body"
/dev_stage="adult"
/lab_host="rabblts"
/clone_lib="sym"
1 213 c 311 g 310
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         US-09-965-594-26 (1-11) x BU718639 (1-1035)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Cyprinus carpio (common carp)
Cyprinus carpio
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Location/Qualifiers
1, 1035
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       311 9
```

9

```
    .502
    /organism-"Danio rerio"
    /mol_type-"mrna"
    /db_xref-"taxon:7955"

                                                                                                                                                                                                                                                                                                                                                                                                         Gaps:
                                                                                                                                                                                                                                                                                                                                                                                                                                                 US-09-965-594-26 (1-11) x BI882963 (1-487)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Danio rerio (zebrafish)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       BE556909.1 GI:9821399
                                                                                                                                                                                                                                                                                                                     41.00
100.00%
63.64%
83.67%
                                                                                                                                                                                                                                                                                                                                       Percent Similarity:
Best Local Similarity:
                                                                                                                                                                                                                                                                                Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                   Query Match:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         DEFINITION
                                                                                                                                                                                                               BASE COUNT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ORGANISM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  TITLE
JOURNAL
                                                                                                                                                                                                                                                                                                      .
9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           RESULT 13
BE556909
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ACCESSION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    REFERENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          AUTHORS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       VERSION
KEYWORDS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   FEATURES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            COMMENT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SOURCE
                                                                                                                                                                                                                                   ORIGIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       LOCUS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           and wilson, A.

Washu Zebrafish EST Project 1998
Unpublished
Contact: Stephen L. Johnson
Contact: Stephen L. Johnson
Washington University School of Medicine.
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Email: zbrafish@watson.wustl.edu
CDNA Library Preparation: Ning Wu. CDNA Library Arrayed by: Steve
CDNA Library Preparation: Ning Wu. CDNA Library Arrayed by: Steve
CDNA Library Preparation: Ning Wu. CDNA Library Arrayed by: Steve
CONTECT Phis clone is available royalty-free through LLNL;
contact the IMAGE Consortium (info.llnl:gov)for further
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      l (bases 1 to 487)
Clark, M. Johnson, J. Lehrach, H., Lee, R., Li, F., Marra, M., Eddy
Clark, M., Johnson, S., L., Lehrach, H., Beck, C., Wylie, T., Underwood
K., Steptoe, M., Theising, B., Allen, M., Bowers, Y., Person, B.,
Swaller, T., Gibbons, M., Pape, D., Harvey, N., Schurk, R., Ritter, E.,
Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann, R., Waterston, R.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    BI882963 487 bp mRNA linear EST 16-SEP-2002 fn01e01.x1 Zebrafish Research Genetics C32 fin Danio rerio cDNA clone IMAGE:4469065 3', mRNA sequence.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Bukaryota, Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes; Cyprinidae; Danio.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               /clone="IMAGE:4469065"
/tissue_type="Fin"
/lab_host="GeneHogs (HS996, a phage-resistant isolate of
                                                                                                     /clone="H7"
/clone_lib="Cyprinus carpio head kidney stimulated by lipo-polysaccharide and concanavalin-A"
/note="common name:common carp; stimulated by lipo-polysaccharide and concanavalin-A"
/note="common name:common carp; stimulated by lipo-polysaccharide and concanavalin-A"
34 c 45 9 70 t
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   /clone_lib="Zebrafish Research Genetics C32 fin"
                                                                                                                                                                                                                                                                                                                   222
7
7
0
0
0
                                                                                                                                                                                                                                                                                                                     Length:
Matches:
Conservative:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   42 GGGCAGTGGTTGTCATGGGAAGAGTTGTACTG 74
                                                                                                                                                                                                                                                                                                                                                                                   Mismatches:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1 GlySerValValIleValGlyArgIleValLeu 11
                                          /organism~"Cyprinus carpio"
/mol_type~"mRNA"
/db_xref~"taxon:7962"
                                                                                                                                                                                                                                                                                                                                                                                                       Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 /organism="Danio rerio"
/mol_type="mRNA"
/db_xref="taxon:7955"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     444.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                   US-09-965-594-26 (1-11) x AUI83349 (1-222)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 High quality sequence stop: 4
Location/Qualifiers
  Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         information.
Seq primer: T7 from Gibco
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    BI882963.
BI882963.1 GI:16090234
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Danio rerio (zebrafish)
                                                                                                                                                                                                                                                                                                                   146
41.00
100.00%
63.64%
83.67%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1. .487
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               DH10B)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Danio rerio
                                                                                                                                                                                                                                                                                                                                                                                     Best Local Similarity:
                                                                                                                                                                                                                                                                                                                                                               Percent Similarity:
                                                                                                                                                                                                                                                                                                 Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              EST
                                                                                                                                                                                                                                                                                                                                                                                                   Query Match:
DB:
                        source
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          DEFINITION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  BI882963/c
                                                                                                                                                                                                                                                                                                                          Pred. No.:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      TITLE
JOURNAL
COMMENT
                                                                                                                                                                                                                                      BASE COUNT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ORGANISM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ACCESSION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 RESULT 12
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        REFERENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          KEYWORDS
SOURCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           FEATURES
FEATURES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          VERSION
```

```
Danio rerio
Eukaryota: Metazoa: Chordata: Craniata: Vertebrata: Euteleostomi:
Actinoptersygil: Neopterygil: Teleostei: Ostariophysi: Cypriniformes
: Cyprinidae: Danio.
1 (bases 1 to 502)
Clark, M. Johnson, S.L., Lehrach, H., Lee, R., Li, F., Marra, M., Eddy
.S. Hillier, L., Kucaba, T., Martin, J., Beck, C., Wylie, T., Underwood
.K., Steptoe, M., Theising, B., Allen, M., Bowers, Y., Person, B.,
Swaller, T., Gibbons, M., Pape, D., Harvey, N., Schurk, R., Ritter, E.,
Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann, R., Waterston, R.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           BE556909 502 bp mRNA linear EST 30-AUG-2000 MRNA sequence. Substance control of the sequence. BE556909
/note-"Vector: pT7T3D-Pac with a modified polylinker;
Site_1: EcoRI; Site_2: NotI; 1st strand cDNA was prepared
from zebrafish(C32) fin, and was then primed with a Not I
oligo(dT) primer. Double-stranded cDNA was ligated to
Eco RI adaptors (Pharmacia), digested with Not I and
cloned into the Not I and Eco RI sites of the modified
pT7T3 vector. Library is non-normalized. Library was
constructed by Ning Wu. NOTE: This clone is available
royalty-free through LLNL; contact the IMAGE Consortium
(info.inl.gov)for further information.

a 115 c 91 g 137 t.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        /clone_llb="Zebrafish Research Genetics C32 fin"
/clone_lbector: pT/T3D-Pac with a modified polylinker;
Site_l: EcoRI; Site_2: NotI; 1st strand cDNA was prepared
from zebrafish(C32) fin, and was then primed with a Not I
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Genome
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  /tissue_type="Fin"
/lab_host="GeneHogs (HS996, a phage-resistant isolate of
DH10B)"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           and Wilson, R.
Washu Zebrafish EST Project 1998
L Unpublished
Contact: Stephen L. Johnson
Washington University School of Medicine
4444 Forest Park Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1810
Email: Zbrafish@watson.wustl.edu
CDNA Library Preparation: Ning Wu. CDNA Library Arrayed by:
Research Genetics. DNA Sequencing by: Washington University Genc
Sequencing Center Clone distribution: Research Genetics web
                                                                                                                                                                                                                                                                                                                                                                                                                                                                     4 B 7 C O O O O
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            address: http://www.researchgenetics.com/
Seq primer: T3 ET from Amersham
High quality sequence stop: 466.
Location/qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Matches:
Conservative:
Mismatches:
Indels:
```

```
Unpublished
                                                                                         Percent Similarity:
Best Local Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                159 a
                            Alignment Scores:
                                                                                                                                 Query Match:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             source
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             BASE COUNT
ORIGIN
                                                                                                                                                                                                                                                                                                                                                                                         DEFINITION
                                                                                                                                                                                                                                                                                                                                                 BI981053/c
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ORGANISM
                                                Pred. No.:
                                                                                                                                                                                                                                                                                                                                                                                                                                                      ACCESSION
VERSION
KEYWORDS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            AUTHORS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        REFERENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     FEATURES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              COMMENT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SOURCE
                                                                        Score:
                                                                                                                                                                                                                                                                                                                             RESULT
                                                                                                                                                                                                                                                                                                                                                                       Locus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Email: thisse@igpmc.ustrasbg.fr
Email: thisse@igpmc.ustrasbg.fr
EST from a cDNA of a gene whose expression is spatially restricted during embrygonessis. We have esstablished its expression pattern during embrygonessis. We have esstablished its expression pattern on zebrafish embryos from the gastrula stage to 2 days of development. The corresponding data are available on the zebrafish community database at http://zfin.org/cDNA library preparation: B. Riggleman. DNA Sequencing by:IGBMC sequencing facility. Clone University of Oregon (Institute of Neuroscience, 1254 University of Oregon. Eugene. OR 97403-1254).
Seq primer: T3 ATTAACCCTCACTAAAGGGA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CB417286 564 bp mRNA linear EST 27-MAR-2003 STR00763 gastrula stage cDNA library Danio rerio cDNA clone CB380 5' similar to CCAAT/enhancer binding protein beta, mRNA sequence.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygil; Neopterygil; Teleostei; Ostariophysi; Cypriniformes; Cyprinidae; Danio.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1, rue Laurent Fries, BP163, CU de Strasbourg, 67404 Illkirch Cedex
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       l (bases 1 to 564)
Loppin, B., Steffan, T., Kempf, J., Heyer, V., Thisse, C. and Thisse, B.
Expression of the zebrafish genome during embryogenesis
Unpublished
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       - oligo(dT) primer. Double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT7T3 vector. Library is non-normalized Library was constructed by Ning Wu. NOTE: This clone is available royalty-free through LLNL: contact the IMAGE Consortium (info.llnl.gov)for further information.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Contact: Thisse B
Institut de Genetique et de Biologie Moleculaire et Cellulaire
CNRS, INSERM, ULP
                                                                                                                                                                                                                                                                               Conservative:
Mismatches:
                                                                                                                                                                                                                                                                                                                                                                                                                                                        355 GGGCAGTGGTTGTCATGGGAAGAGTTGTACTG 387
                                                                                                                                                                                                                                                                                                                                                                                                                                  1 GlySerValValIleValGlyArgIleValLeu 11
                                                                                                                                                                                                                                          Length:
Matches:
                                                                                                                                                                                                                                                                                                                             Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     /organism="Danio rerio"
/mol_type="mRNA"
/db_xref="taxon:7955"
                                                                                                                                                                                                                                                                                                                                                                                         US-09-965-594-26 (1-11) x BE556909 (1-502)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CB417286.1 GI:29294527
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Danio rerio (zebrafish)
Danio rerio
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Tel: 33 3 88 65 33 60
Fax: 33 3 88 65 32 01
                                                                                                                                                                                                                                                                               100.00%
63.64%
83.67%
10
                                                                                                                                                                                                                                          434
                                                                                                                                                                                                                                                                                                          Best Local Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CB417286
                                                                                                                                                                                                                                                                                    Percent Similarity:
                                                                                                                                                                                                                      Alignment Scores:
                                                                                                                                                                                                                                                                                                                           Query Match:
DB:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     source
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               BASE COUNT
ORIGIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             LOCUS
DEFINITION
                                                                                                                                                                                                                                              Pred. No.:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ORGANISM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          AUTHORS
TITLE
JOURNAL
COMMENT
                                                                                                                                                       SASE COUNT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       VERSION
KEYWORDS
SOURCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ACCESSION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          REFERENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   RESULT 14
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             FEATURES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CB417286
                                                                                                                                                                                                                                                                                                                                                                                                                                  ð
```

```
Contact: Stephen L. Johnson
Mashington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
7444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
7442 86 1800
Fax: 314 286 1810
Email: zbrafishtewatson.wustl.edu
cDNA Library Preparation: John Ngai. cDNA Library Arrayed by:
Matthew Clark. DNA Sequencing by: Mashington University Genome
Sequencing Center Clone distribution: Genome Systems, St. Louis,
Missoult (web address: www.genomesystems.com) (email contact:
info@genomesystems.com) and Research Genetics, Huntsville, Alabama
(web address: www.resgen.com) (email contact: info@resgen.com) and
RessourcenZentrumPrimarDatenbank, Berlin, Germany (web address:
                                                                                                                                                                                                                                                                                                                                                                                                                                   fu38h04.x1 zebrafish adult brain Danio rerio cDNA clone
IMAGE:5332062 3' similar to TR:097894 097894 CCAAT/ENHANCER BINDING
PROTEIN BETA;, mRNA sequence.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes
; Cyprinidae; Danio.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1 (bases 1 to 584)
Clark, M. Johnson, S.L., Lehrach, H., Lee, R., Li, F., Marra, M., Eddy
S., Hillier, L., Kucaba, T., Martin, J., Beck, C., Wylle, T., Underwoo,
K., Steptoe, W., Thelsing, B., Allen, M., Bowers, Y., Person, B.,
Swaller, T., Gibbons, M., Pape, D., Harvey, N., Schurk, R., Ritter, E.,
Schu, S., Shin, T., Jackson, Y., Cardenas, M., McCann, R., Waterston, R.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   /dev_stage_"adult".
/lab_host="E. coli DH10B"
/lab_host="E. coli DH10B"
/clone_lib="scbrafish adult brain"
/clote="vector: pZIPLOX; Site_1: Not1; Site_2: Sal1;
Original library was constructed in lambdaZIPLOX. M
excision of the cDNA library was performed to yield
pZIPLOX plasmids. Insert check was done in original
library.
a 151 c 118 g 156 t
564
7 4
0 0
                                                                 Conservative:
                                                                                                                                                                                                                                                                                                       Mismatches:
Indels:
                                                                                                                                                                                                                                                                        1 GlySerValValIleValGlyArgIleValLeu 11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              /sex="mixed male and female"
/tissue_type="brain"
                               Matches:
      Length:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      and Wilson, R.
WashU Zebrafish EST Project 1998
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Seg primer: -400P
High quality sequence stop: 422.
Location/Qualiflers
1. 584
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    /organism="Danio rerio"
/mol_type="mRNA"
/db_xref="taxon:7955"
/clone="IMAGE:5332062"
                                                                                                                                                                                                             JS-09-965-594-26 (1-11) x CB417286 (1-564)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         BI981053
BI981053.1 GI:16369292
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Danio rerio (zebrafish)
Danio rerio
                           41.00
100.008
63.648
83.678
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         www.rzpd.de)
```

	Length:	Matches:	Conservative:	Mismatches:	Indels:	Gaps:
	531	41.00			83.678	12
Alignment Scores:	Pred. No.:	Score:	Percent Similarity:	Best Local Similarity	Query Match:	DB:

US-09-965-594-26 (1-11) x BI981053 (1-584)

oy ob

12

Search completed: August 31, 2003, 04:27:56 Job time: 112.667 secs